DOCUMENT RESUME

ED 346 105 TM 017 189

TITLE	FairTest Examiner, Volumes 1-5, 1987-1991.
INSTITUTION	National Center for Fair and Open Testing (FairTest), Cambridge, MA.
SPONS AGENCY	Ford Foundation, New York, N.Y.; Rockefeller Family Fund, Inc., New York, N.Y.
PUB DATE	91
NOTE	319p.
AVAILABLE FROM	FairTest, Box 1272, Harvard Sq. Station, Cambridge, MA 02238 ($15.00 per year for individuals and $25.00 per year for institutions).
PUB TYPE	Collected Works - Serials (022)
JOURNAL CIT	FairTest Examiner; v1-5 1987-1991
EDRS PRICE	MF01/PC13 Plus Postage.
DESCRIPTORS	Achievement Tests; Educational Change; *Educational Testing; Elementary Secondary Education; Higher Education; Newsletters; Racial Bias; Sex Bias; *Standardized Tests; *Test Bias; *Testing Problems; *Test Use
IDENTIFIERS	*National Center for Fair and Open Testing; *Reform Efforts; Teacher Competency Testing

ABSTRACT

The FairTest Examiner is a quarterly newsletter that discusses the appropriate use of tests and test results. Many of the features address recent events in the field of educational testing reform. The bulletin's regular features include: current news items, a list of recommended reading material, a listing of FairTest personnel, announcements of speaking engagements by FairTest staff members, a FairTest calendar of events, a listing of FairTest publications, and an order form for obtaining selected publications. The bulletin often includes a short test and suggested answers which highlight problematic items on widely used standardized tests (LFD).

****
*   Reproductions supplied by EDRS are the best that can be made   *
*   from the original document.   *
****

ED346105

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it

☐ Minor changes have been made to improve reproduction quality

- Points of view or opinions stated in this document do not necessarily represent official OERI position or policy

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

CINTHIA SCHUMAN

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

THE FAIRTEST EXAMINER

VOLUMES 1-5, 1987-1991

TM017189

BEST COPY AVAILABLE

the
FairTest Examiner

Volume 1, No. 1 Spring 1987

A publication of the National Center for Fair & Open Testing

"60 years of idiocy is enough"

-- Ralph Nader

"Standardized tests are a specialized form of fraud," consumer rights advocate Ralph Nader told testing reform activists at FairTest's annual convention.

"Multiple-choice tests influence much more than just who gets into colleges, graduate schools, or is able to work in over 100 professions," Mr. Nader explained. "Our reliance on these exams is leading America towards standardized minds, standardized thinking, a standardized curriculum. Any system which controls the standards by which we are judged is a system that controls us.

"Tests have proliferated partly because of the testing industries' promotional practices and partly because the people in charge find it easiest to judge people numerically and not qualitatively. As a result, schools and colleges compete with one another to boost test scores. Curriculum is blatantly shaped and distorted to conform to the limited content of what a standardized multiple-choice test measures."

What is being tested?

"These tests ignore determination, idealism, wisdom, strategic reasoning, judgement, experience, persistence, stamina, creativity, and often writing ability. Do you ever hear of school systems being evaluated on the quality of their libraries? By the kind of citizen training they give their students? What in the world are we doing ignoring 99% of the really important factors in evaluating educational performance and service?

"Our mission should not be to simply make fairer multiple-choice exams or to use Truth-in-Testing laws to expose the wretched quality of test questions. Such reforms are needed to get the foot in the door, but our national obsession with standardized tests demands much more.

"Multiple-choice tests militate against intellectual nuances, against learning how to write the English language, against the creativity of the human mind--which is not standardized. It is absurd to believe that human potential can be judged by a multiple-choice test taken under time pressure."

Self-inflicted wound

"From an important perspective, ETS is one of the most powerful corporations in the world. What other industry is so intimately involved in shaping the careers and self-perceptions of millions of peo-

(continued on page three)

"Curriculum is blatantly shaped and distorted to conform to the limited content of what a standardized multiple-choice test measures."

SPECIAL CONFERENCE ISSUE

This first issue of the *FairTest Examiner* reports on our annual convention, ***"Testing Reform Initiatives: 1987 & Beyond"***, held in Washington, D.C. on December 12, 1986. Educators, legislators, students and civil rights activists from more than 20 states gathered on Capitol Hill to hear keynote presentations by Ralph Nader and Eleanor Smeal, and to help FairTest plan its future agenda.

Contents, page 2

TM017189

SAT: Why are women losing out?

Representatives from major feminist organizations, including the National Organization for Women (NOW), the Project on Equal Education Rights (PEER), Women in Science, the American Association of University Women, and Wider Opportunities for Women joined testing reform activists at FairTest's December convention to address the devastating implications of sex-bias in testing.

No SATisfaction

Much of the discussion revolved around sex bias on the Scholastic Aptitude Test (SAT). ETS' sole justification for the SAT is that it serves as a predictor of first-year college grades. Yet, as education researcher Phyllis Rosser explained, "Girls have higher grades than boys in all subjects in high school and college, even in freshman year, but they averaged 61 points lower than boys on the SAT last year. A test taken by 800,000 females every year that doesn't accurately predict their abilities should be illegal. If boys were getting higher grades and lower test scores, the test would be rewritten."

Rosser added, "Girls are receiving lower scores on the American College Test (ACT) as well, averaging 1.5 points lower than boys (on a 36-point scale). Unlike the SAT, where girls score lower on both the Math and Verbal sections, girls taking the ACT score slightly higher in English, although they remain much lower in Math."

Double Jeopardy

"The SAT places minority women in double jeopardy," explained Northeastern University Law Professor and Chair of FairTest's Board, Denise Carty-Bennia. "Not only do they suffer the racial and cultural biases in these exams, but they then face the additional hurdle of sex bias, truly eliminating them from competition."

The negative effect of low admissions test scores manifests itself in number of ways. Girls may be kept out of certain college programs which rely heavily on scores from standardized admissions tests. Lower test scores also reduce girls' chances of obtaining scholarships based on 'merit'. For example, qualification for National Merit scholarships is based on scores of ETS' Preliminary Scholastic Aptitude Test (PSAT). Girls scored an average of 53 points lower than boys on this test last year. In some areas, such as West Hartford, Connecticut, only 20% of last year's National Merit semifinalists were female. This poor showing is replicated nationally as approximately twice as many boys as girls are eligible for the $30 million worth of National Merit scholarships awarded annually (see chart on page 3).

Perhaps the most devastating effect of unfairly low test scores is the way they undermine girls' perceptions of their own abilities. "I've learned in fighting discrimination that one of the most difficult parts is that we, who are the victims, accept the standards," said NOW President Eleanor Smeal. "We put up a barrier, and think we can't walk through it. If we can't walk through it, eventually it becomes a self-fulfilling prophecy."

Sue Klein, a senior research associate at the U.S. Department of Education, recommended that test publishers divulge statistics relating to the national pattern of test results by gender, race, ethnicity, and socioeconomic status. "I was surprised to find that test publishers have this information on their computer screens, but they don't analyze it and share it with the public."

1986 National Merit Scholarship qualifiers by sex

State	Boys	Girls
California	817	493
Idaho	49	14
Massachusetts	305	178
New Jersey	324	173
New Mexico	66	24

CONTENTS

- Sex Bias in Testing page 2
- The Golden Rule Bias Reduction Procedure page 4
- Victory in Alabama: Teacher Testing page 5
- "How I Fought ETS...and Lost" page 6
- An Open Letter to the College Board page 7
- K-12 Tests: Impact on Access & Excellence page 8
- Victory in California: IQ Testing page 10
- FairTest's Campaign for an Independent ERIC page 11
- FairTest Job Openings, Activities Update page 14
- Can You Pass This Test? page 15
- Subscription Form back cover

The *FairTest Examiner* is a publication of the National Center for Fair & Open Testing (FairTest), P.O. Box 1272, Harvard Square Station, Cambridge, MA 02238, phone (617) 864-4810≥ We are a national research and advocacy organization working to ensure that the 40-plus million standardized multiple-choice exams administered annually to America's students and job applicants are fair, open and educationally-sound.

We would like to thank the Rockefeller Family Fund for partially funding the start-up costs of this publication.

Late Breaking News

A March 1987 research report by FairTest and the New York Public Interest Research Group (NYPIRG) has revealed that more than twice as many high school boys as girls were awarded the prestigious New York State-funded Empire State scholarships in 1987. Recipients were selected by their SAT & ACT scores. One thousand scholarships--worth up to $10,000 each--were awarded in 1987.

Average SAT Scores

	Females	Males	Diff.
Asian-Pacific Americans	897	946	-49
Black	705	748	-43
Mexican-American	775	845	-70
Native Americans	790	855	-65
Puerto Rican	744	820	-76
White	912	969	-57
National Average	877	938	-61

--from 1985 Profiles, College Bound Seniors, *by Leonard Ramist and Solomon Arbeiter, CEEB 1986.*

N.O.W. to Fight Testing Bias

"The first part of fighting bias is to expose it," National Organization for Women President Eleanor Smeal told the plenary session at FairTest's December convention. "When I lecture on America's college campuses, young women's faces light up when I describe how the SAT systematically underpredicts the ability of females. Just telling people about sex bias on these exams is a big liberator."

Differences in test scores translate into dollars for scholarships and admissions to colleges, the tickets to higher paying jobs, Smeal said. Unfairly low test scores often become a "self-fulfilling prophecy...The SAT helps destroy girls' self-confidence. They internalize it, and feel that they cannot compete."

Smeal concluded her talk by calling for a multi-pronged approach to do away with biased testing. "NOW is investigating the possibility of using Title IX litigation to prevent university financial aid offices from basing their decisions on SAT scores." She also urged that the issue be introduced into women's studies programs nationwide, both to increase public awareness and to generate interest in further research.

"Just telling people about sex bias on these exams is a big liberator."
--Eleanor Smeal, NOW President

Bibliography available

For a bibliography on sex bias in aptitude and achievement testing, send FairTest a stamped, self-addressed envelope.

April 16, 1987
Seminar on Sex Bias

"Strategies to Eliminate Sex Bias from Tests" will be the theme of an invitational seminar at Hunter College in New York City. For more information, contact FairTest (617) 864-4810.

RALPH NADER

(from page 1)

ple each year? Their pseudoscientific tests not only shape what is taught in our schools, but penetrate the minds of test-takers who internalize into their own psyches ETS' verdict of their abilities.

"Standardized testing is one of the greatest self-inflicted wounds...ever imposed on the American people. It has had a life of about sixty years. Sixty years of idiocy is enough!"

Looking ahead, Nader demanded that test reformers seek ways to raise the temperature of their cam paigns. "We're not dealing with anything minor here. We're dealing with nothing less than the liberation of human potential, of the human capacity for diversity. Students must pressure parents, educators and policy-makers to bring about a fundamental change. Only in this way can we defeat the kind of serfdom which these tests represent."

EXTENDING GOLDEN RULE:

The campaign intensifies

For decades, civil rights leaders and educational reformers have claimed that racial and gender bias pervades standardized, multiple-choice exams. But without an objective technique to detect and replace unfair items, test critics have been unable to force significant changes.

The landmark agreement between the Golden Rule Insurance Company, the Educational Testing Service and the State of Illinois in November 1984 forever altered the testing landscape. The settlement of a decade-old lawsuit charging ETS' Illinois Insurance Agent Licensing Exam unfairly discriminated against blacks and was not job-related gave reformers a simple tool to make exams fairer, but not easier.

Under Golden Rule, the same content areas are covered as on previous tests, and the exams are of the same overall level of difficulty. The only difference is that within groups of equally-difficult items in the same content area, test publishers must select those items which display the least difference in correct answer rates between majority and minority test-takers. Thus, Golden Rule assures that exams measure real knowledge differences among test-takers and not irrelevant, culturally-specific information.

The Illinois Experience

Data presented to the Advisory Committee established to monitor implementation of the Illinois settlement reveals the significant impact of the Golden Rule reforms. In the first round of exams assembled using the new procedure, the difference between black and white test scores was reduced by one quarter. Since many blacks previously fell just below the passing level "cut score", the percentage of minorities who become insurance agents should increase significantly. Experts believe the test score gap will narrow even further as future editions of the exam are compiled using the Golden Rule bias reduction technique.

Golden Rule in the States

At FairTest's Capitol Hill conference, activists from across the country described campaigns to extend Golden Rule. In addition to implementing the landmark victory in Alabama's Teachers Licensing Test, 1987 will see legislative initiatives in at least half a dozen other states.

- In **New York State**, a coalition of civil rights, feminist, consumer and student organizations led by the New York Public Interest Research Group (NYPIRG) is launching a major drive to enact legislation establishing a state commission to conduct an item-analysis of college and graduate school admissions exams. The commission would use simple statistical techniques to examine the scoring patterns of white, minority, male and female test-takers and suggest ways to reduce unexplained discrepancies. Chief legislative sponsors Sen. Kenneth LaValle (R.--Port Jefferson) and Assembly Deputy Speaker Arthur Eve (D.--Buffalo) expect to build on last year's successes, in which the same coalition won adoption of three proposals extending New York's landmark Truth-in-Testing law (see *FairTest News Update* #4). The group will also file bills to extend Golden Rule and Truth-in-Testing protections to the state's employment licensing exams.

- In **Texas**, the Mexican American Legal Defense and Education Fund (MALDEF) is drafting legislation to apply the Golden Rule Principle to the state's teacher licensing exams. The effort complements an ongoing lawsuit claiming ETS' Pre-Professional Skills Test (PPST), used for admission into teacher training programs, is racially biased.

- Golden Rule extension bills will also likely be filed in **California, Massachusetts, Rhode Island** and **Wisconsin**.

1987 GR Sourcebook

An updated 1987 edition of FairTest's *Golden Rule Bias Reduction Procedure Sourcebook* is now available. This 100-page manual includes copies of model legislation, fact sheets, resource lists, the response to Golden Rule by ETS and other test companies, and an analysis of test companies' false and misleading statements about FairTest's initiatives. Available for $20 from FairTest.

We understand you haven't yet subscribed to the FairTest Examiner*! (see back cover)*

Teacher Testing: Victory in Alabama

In 1981, four black teachers and a predominantly black state university sued the state of Alabama after a disproportionate number of minorities failed the state's teacher-certification exam constructed by National Evaluation Systems of Amherst, MA. The plaintiffs in *Margaret T. Allen v. The Alabama Board of Education* claimed the tests discriminated against blacks, violating their 14th Amendment protections and federal civil rights laws.

The stormy five-year history of the case was described by the plaintiff's lawyer, Don Watkins, at the FairTest conference. After hearing testimony on the test's validity from expert witnesses, including former President of the National Council of Measurement in Education Professor George Madaus, an out-of-court agreement was reached between the teachers and the State Board of Education.

Golden Rule Concept Extended

Under the terms of the settlement, new tests were to be created by using a modified version of the Golden Rule procedure. Furthermore, the test was eliminated as the sole criterion for certification. Instead, a sliding scale including grade point averages was mandated in cases where candidates failed the exam. The plaintiff's lawyers were also given the right to monitor the settlement to make certain it would be fully carried out. Finally, the settlement prohibits testing on material not taught in the teacher education curriculum and establishes separate review panels with minority members to examine the tests for cultural and racial bias.

Unfortunately, Alabama's Attorney General seized on the settlement as an issue in his campaign for higher office. The public confusion caused the state Department of Education to vote to rescind its original agreement. But on November 25, 1986, the U.S. Court of Appeals reinstated the settlement, ordered the state to re-vamp its teacher certification testing program, and to pay $500,000 in damages to those who failed the test. Subsequently, the Board of Education voted to appeal the decision. But the settlement remains in effect while legal maneuvering continues.

'A testing travesty'

Attorney Don Watkins, a self-described 'country lawyer', gave FairTest conference participants a vivid description of how he came to the testing field:

"I didn't know about test-making. It sounded so scientific when I got started." But then he took a closer look at the test made by National Evaluation Systems and discovered that the first version had been failed by every one of the first 800 people who took the test, black and white.

"Once I stripped away the fancy covers on the test booklet, once I kicked aside the reams of computer print-out, once I got past the test-makers' promotional brochures, I realized that this country lawyer had nothing more than another product to look at, and to be evaluated in the same way as any other product involved in a lawsuit. I looked at it. I needed to find out: Is this product built well? Is this product hurting people? What is the reputation of the person or group who makes the product? What research and development went into this product?"

Watkins concluded that the product was a travesty--shoddily constructed and never properly validated. In court, he argued the test's manufacturer had promised the State of Alabama something which it never made a serious effort to deliver. After four years and hundreds of thousands of hours and dollars, Watkins and his team helped free Alabama's teachers from "a defective, worthless product that is hurting people who deserve better from their public policy-makers."

"Nobody is watching these people. School boards don't know what the hell they are getting." -- Don Watkins

Adventures in admissions testing: How I fought ETS and lost

The following article has been taken from the Chronicle of Higher Education, *January 14, 1987.*

By John L. Grigsby

THE QUESTIONABLE INTEGRITY of the Educational Testing Service and the unfairness and inaccuracy of its tests are among the most durable subjects of debate in academe.

However, E.T.S. sees all criticism of its products as an attack on its profitable business, so any criticism, however accurate it may be, is ignored or, if that is impossible, resisted vehemently. The profitability of the tests is such that valid accusations of unfairness, inaccuracy, and even injustice have no effect on the way they are set up or administered, unless the force of law is behind the accuser. New York's "truth in testing" law was a step in the right direction, but E.T.S. needs to be made even more accountable to the people who must take its tests.

My experience illustrates this point.

In the fall of 1972, when I was a senior at a small private college in Kentucky, I registered to take the Graduate Record Examination. Although I knew G.R.E. scores were generally required for admission to graduate school, I didn't prepare in any special way for the test. I had very good grades and very good recommendations and assumed that those would be much more important to admissions officers than scores on a standardized test.

Indeed, I even forgot about the G.R.E., until one Friday night (actually about 4 o'clock one Saturday morning) weeks later when, in the midst of a continuing T.G.I.F. party, I heard from a friend that the test was scheduled for that very Saturday. With great effort, I pulled myself together, staggered back to the dormitory, and fell into bed. Three hours later, the alarm woke me and, with a splitting headache and a very upset stomach, I dragged myself off to take the test.

Actually, the G.R.E. took me. When I got my scores back they were not good, but I did manage to get admitted to the least prestigious of the several M.A. programs to which I had applied. I realized then that if I wanted to get into the kind of Ph.D. program I had in mind, I would have to improve on my poor performance on the Graduate Record Examination.

THE Educational Testing Service's brochure for the G.R.E. states that anyone who is dissatisfied with the results is free to retake the test. So, after a semester in graduate school, I applied to Ph.D. programs at 10 well-known institutions and registered to take the exam again.

This time I prepared carefully, got plenty of sleep instead of partying the night before, and came to the testing site alert and confident. I left it feeling I had done very well and that admission to the kind of program I aspired to was now a real possibility.

Not long after, however, I received a short, somber letter from E.T.S. saying that I had scored so much higher on the retake than I had the first time that the retake was being investigated. I was called in and questioned by the person who had administered the test. I was honest—still am, in fact—about my condition on the first testing date and the reason for it, and explained that I had prepared much more wisely the second time. The test administrator was sympathetic, and I left his office encouraged. Hearing nothing further from E.T.S., I was sure my retake scores would be accepted.

> "Its power to determine the course of so many students' careers should be reason enough for stricter laws to control E.T.S."

When I was turned down by the Ph.D. programs that were my first three choices and also by five of the other seven programs I had applied to, I was philosophical. Maybe I had been right the first time in thinking the G.R.E. scores weren't very important, and my grades and recommendations just hadn't been good enough.

A year or so ago, older and wiser after earning my Ph.D. and teaching for several years, I began to think about what happened. It occurred to me that I might well have been "had" by the E.T.S. With the experience and wisdom of middle age, I reasoned that, given the quality of my grades and recommendations, with good scores on the G.R.E. I should have been admitted to at least one of my top choices.

With that in mind, I wrote to E.T.S. and asked for my G.R.E. scores. When they arrived several weeks later, I was dismayed to see that only those on the first test were listed. I called E.T.S. and was bluntly told that my retake scores did not exist, that they had been deliberately destroyed. Justifiably angry now, I demanded to know why. I was told that it must have been because the investigators had decided I had cheated on the retake. Apparently the odds (as calculated by E.T.S.) against that much improvement between takes seemed so great to the investigators that they simply assumed I had somehow broken test security and had seen the questions beforehand.

KEEPING THE TESTS SECRET, of course, makes for a highly profitable operation, since few changes have to be made in the questions (and few improvements, too). E.T.S.'s paranoia about test security was behind its desperate resistance to New York's truth-in-testing law, and was also the reason it had punished me years before for legitimately improving my scores. Despite my plausible and truthful explanation, my second scores were thrown out without my even being notified. Hence, by duplicity, E.T.S. essentially prevented me from taking the exam again and, perhaps, from gaining admission to a more prestigious Ph.D. program. Graduating from such a program would, of course, have increased my chances of career success.

E.T.S. is so concerned with maintaining the profitable status quo that it ignores the fact that it is adversely affecting the future of people who must take the tests but who do not conform to its cultural bias or, like me, do not fit its mathematical-probability calculations. Its power to determine the course of so many students' careers should be reason enough for more and stricter laws to control E.T.S. In the meantime, it would help if colleges stopped using the tests altogether.

John L. Grigsby is assistant professor of English at Mississippi Valley State University.

FairTest-Stewart meeting

An Open Letter to CEEB

Immediately following President-Elect Donald Stewart's inaugural speech to College Board delegates at their annual convention, FairTest sent Dr. Stewart an open letter urging him to implement "11 practical, inexpensive steps to improve the quality of the Scholastic Aptitude Test". Students distributed copies of the 10-page letter to the 2,000 high school counselors and admissions officials attending the convention at New York's Waldorf-Astoria Hotel. Dr. Stewart has agreed to meet FairTest representatives.

FairTest recommends:

1) **Colleges should independently evaluate their need for the SAT.** The College Board should discourage colleges from using validity studies generated by the Educational Testing Service (ETS) and, instead, independently investigate their own need for the tests. Relying on the recommendation of ETS, which produces the $50 million-a-year SAT package, is like asking General Motors to evaluate what brand of car you should buy.

2) **Open up the SAT for competitive bid.** The College Board should put the contract to develop its admissions testing program out for public bid. For the past 40 years, the College Board has simply renewed its agreement with the Educational Testing Service. Internal documents reveal that ETS earns a profit of about 30% from its College Board-related products.

3) **Make the test as fair as possible.** The SAT should be modified so that it no longer underpredicts the ability of high school girls (see pages 2-3). The Golden Rule Bias Reduction safeguards should be applied to the SAT (see page 4).

4) **Withhold SAT scores from institutions which misuse them.** The College Board must take responsibility for the use of its tests by refusing to transmit scores to colleges and scholarship agencies which misuse them. The American Association of Collegiate Registrars & Admissions Officers found more than 40% of colleges automatically denied admission to students who failed to achieve an arbitrary test score, a direct violation of current College Board guidelines.

5) **Require colleges, not students, to pay for the test.** The cost of the test should be paid by those who use it: admissions officers. If colleges were billed for the SAT, they would have to conduct cost-benefit analyses of their need for the exam. At present, many schools which admit nearly all applicants require the test solely to maintain an aura of selectivity.

6) **Rename the SAT.** Many individuals and institutions place too much weight on SAT scores in part because of the misleading word "Aptitude" in its title. But neither the College Board nor ETS is able to define "Aptitude" or prove the SAT measures it. Since the only statistical justification for the test is its ability to predict Freshman grades, "The Test that Attempts to Predict First Year College Grades" would be a more accurate title.

7) **Add a warning to test score "wall charts".** Anyone who attempts to rate educational systems by test scores should be required to include a warning consistent with College Board Senior Vice President Daniel Taylor's testimony before Congress that "group average scores on the Scholastic Aptitude Test...are not an appropriate or scientifically valid basis for comparing the overall performance of teachers, schools, districts, states or other groupings..."

8) **Study whether coaching actually improves SAT scores.** The College Board should cooperate with FairTest in a study to clarify once and for all the effectiveness of test preparation courses. The College Board and ETS claim that coaching doesn't work, yet last year students and their parents spent well over $50 million on such courses.

9) **Inform students about the experimental section.** Before administering the experimental section of the SAT which contains especially difficult and ambiguous questions used solely for data collection, the College Board should obtain the informed, written consent of each student. This would be consistent with ethical standards for experimentation on human subjects.

10) **Extend New York's 1986 Truth-in-Testing Laws Nationwide.** The College Board should voluntarily extend New York's new laws protecting test takers to all 50 states. The provisions require test companies to publicly defend challenged questions, allow students to request copies of their scored answer sheets when they sign up for exams, and open the Test of Standard Written English to public inspection.

Note: For a copy of FairTest's complete letter, send a self-addressed envelope with a 39¢ stamp to FairTest.

Brown SAT referendum

"Should Brown University continue to use the SAT in its admissions decisions?" This referendum question will appear on the ballot in April's student elections.

The K-12 tests: Impact on Access & Excellence

-- *PTA President Ann Kahn, National Urban Coalition Education Director DeAnna Beane, NAACP Education Director Beverly Cole, and ACRE Director Page McCullough.*

Should five-year-olds be routinely administered standardized tests which label half of them "worse than average" before they even enter first grade? Should one test score be allowed to determine a child's entire future? Do state-mandated standardized tests measure the critical thinking and analytical skills needed in the 21st century?

These were some of the questions raised by FairTest Conference participants who discussed issues related to state-mandated K-12 testing amid the martial art of the Rayburn Building's armed forces subcommittee offices. The huge murals of regimented ranks of soldiers seemed curiously in keeping with the stories of test-driven conformity in the nation's schools.

Test-driven instruction

PTA President Ann Kahn, National Urban Coalition Education Director DeAnna Beane, NAACP Education Director Beverly Cole, and Page McCullough of the Atlantic Center of Research in Education (ACRE) all vividly described how state-mandated testing is contributing to a rigidly-tracked school system which narrows the curriculum and limits options for disadvantaged students. They pointed to the paradox that these exams had become an end in themselves, with school curricula debased and students often bored, humiliated, and internalizing a sense of personal failure.

Speakers lamented the fact that tests were not being used for diagnostic purposes, but as barriers, and inculcated a dangerous 'blame the victim' syndrome.

"The tests are redundant," PTA President Ann Kahn told participants. "We test when we already know the answer." Everyone should ask the right question: "Who needs the information from these tests? What is it for?" Ms. Kahn said the multiple-choice test mentality is ingrained long before high school. She argued that the enormous pressure to teach 'basics' had terrible effects on the school curriculum, with students finding little incentive to learn writing skills and having less and less understanding of how to apply the 'basics' to real-life situations. She concluded that "the fabulous amounts of money spent on testing might be better spent on services for children or opportunities to improve teaching."

The PTA's position is that tests should not be given if they were not directly related to what was being taught in the classroom. On the other hand, the PTA recognizes that it will be difficult to get rid of standardized tests unless parents can be educated about what they actually mean, and reasonable alternative forms of assessment developed.

"Stranglehold"

Page McCullough described how similar concerns led to the founding of ACRE in 1978 and how state-mandated tests had shifted power away from local school districts to the state capitol. She maintained that the tests were a stranglehold on the education system; the extension of a class system which uses 'objective' numbers to disguise the way we decide who is to get the rewards. But Ms. McCullough also recognized that, for many minority parents, standardized tests are fairer than some teachers ever were, and that low test scores could be used by parents to put pressure on the schools to do more. She encouraged working for reforms, including a testing bill of rights which would ensure that exam scores are not used for tracking purposes and extensions of both Truth-in-Testing and Golden Rule procedures.

Looking ahead

Beverly Cole of the NAACP painted a bleak picture of what the use of state-mandated tests to track students meant for black children, and for the future of the country. Citing the fact that minority groups would be more than one-third of the country's population by the year 2000, she predicted that "a large proportion of American society is not going to be educated." If certification exams remain, soon there will be very few black teachers. Test scores, she observed, put pressure on everyone concerned. They lead administrators to juggle scores to make their schools look good and they are internalized by students. Yet many students are tested on subjects such as algebra and geometry even though they have never taken courses in those areas.

Coaching for survival

But as long as the tests remain in place, the NAACP believes that children must be offered instruction in test-taking skills so they can be promoted, graduate, and get into college. For this reason, the NAACP sponsored a coaching course which resulted in a 100-point increase for many students on the SAT. But at the same time, the NAACP believes tests must be demystified, for both students and their parents. And parents have to be re-educated to understand that mandating another battery of tests does not necessarily have anything to do with meaningful education reform.

Diagnostic tests only

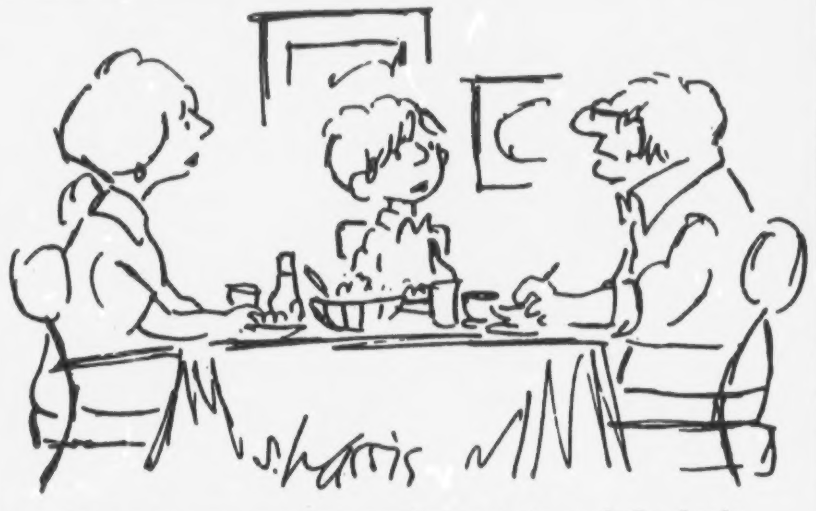
The National Urban Coalition's DeAnne Beane was appalled at the growth of the minimum competency movement. In many states, these tests are given at each grade level, and in 18 states they are required for high school graduation. This is leading to a 'dual diploma' system which has far-reaching implications. She wondered what the effect of the minimum competency movement was on the high school dropout rate, which in New York City was already approaching the staggering figure of 70% for black males. If we are stuck with tests, Ms. Beane concluded, we should work to change their nature and the way they are used, insisting they only be given for diagnostic purposes. Parents must be empowered to help bring about this change.

Other workshop participants took up several of the points made by the presenters. Woody Kelley of the Department of Government at Cornell said that change should be sought at the level of local school districts; they should accept the results of state-mandated tests only to the extent that those tests reflected the school curriculum. School districts should also be made to insist that no child be placed in a different track because of a test score, unless that score moved him or her into a higher one.

> **"By 2000...a large proportion of American society is not going to be educated."**
> ***-- Beverly Cole,***
> ***Education Director, NAACP***

> **"The fabulous amounts of money spent on testing might be better spent on services for children or opportunties to improve teaching."**
> ***-- Ann Kahn, President,***
> ***National PTA***

Lesley College's George Hein challenged the testing industry's assumption that children can be rank ordered along a bell-shaped curve of normal distribution. Dr. Hein argued that there was "nothing normal about it," and urged conference participants to publicize the numerous educationally-sound, non-multiple-choice methods of evaluating children.



"What with the reading readiness tests and the basic skills tests and the I.Q. exams and the sequential tests of educational progress and the mental maturity tests and the minimum competency exams and the grade promotion tests and my prep course for the PSAT, we haven't been learning anything at school."

The Larry P. Settlement

I.Q. tests banned in California

On September 25, 1986, Federal District Court Judge Robert Peckham signed an order stipulating that IQ tests may not be administered to black children in California *for any purpose whatsoever*, even when parents have given their permission. The state has decided not to appeal this ruling, so after 15 years of litigation, the *Larry P. vs. Riles* lawsuit is finally over.

"The facts are there, the precedent has been set," Psychologist Harold Dent, chair of the Testing Committee Association of Black Psychologists told FairTest Conference participants. "We hope to encourage more such lawsuits."

According to San Francisco State University Professor Gerald West, who testified at the trial, "The *Larry P.* case is significant because it is the first successfully-litigated case disproving theories of racial intellectual inferiority."

Long Journey of Larry P.

In 1971, a group of black parents in San Francisco filed a class action lawsuit charging that the WISC-R, Stanford-Binet and other IQ tests were culturally-biased and should not be used to track black children into "dead end" Educably Mentally Retarded (EMR) classes. The *Larry P.* trial began in 1977, with Judge Peckham ruling in 1979 that these IQ tests were culturally biased, and had never been validated for black children. The judge ordered San Francisco Public Schools to stop using IQ tests to place black children in EMR classes.

After a lengthy appeal, the District Court's findings were upheld. But in the early 1980s, psychologists and lawyers monitoring the settlement found that, despite the Court's ruling, schools were still misusing IQ tests to mislabel black children. Last November, after hearing additional arguments, Judge Peckham banned the administration of IQ tests for black students for any purpose whatsoever. The state has voluntarily extended the ruling to all public schools in California.

"We hope to encourage more such lawsuits."

- Harold Dent

ARE I.Q. TESTS FAIR?

The following two questions are taken from Harcourt-Brace's Wechsler Scale for Children Revised (WISC-R), which is currently the largest-selling IQ test in the country. The WISC-R and all other IQ tests can no longer be administered to black school children in California.

1) **What is the thing to do when you cut your finger?**
   2 point response: Put a bandaid on it
   1 point response: Go to the hospital
   0 point response: Cry, bleed or suck on it

--WISC-R IQ Test

Minority children usually perform poorly on this item. A few years ago a Baltimore, Maryland sociologist asked several inner-city youths why they answered the question the way they did. She found that many of these kids answered "go to a hospital" because they thought that cut meant a big cut. Or they thought it was a small cut--and since they didn't have any bandaids at home, they answered "sucked on it"--and received no points.

2) **What is the thing to do if you lose a ball that belongs to one of your friends?**
   2 point response: Buy him a new one and pay for it
   1 point response: Look all over for it, try to find it
   0 point response: I'd just cry, tell him you're sorry, apologize

--WISC-R IQ Test

If you're poor, it often is not possible to buy another one and it is unlikely that you have one to give. If the ball is really lost, looking for it is a waste of time. And if it is not really lost but could be found, wouldn't looking for it be more 'intelligent' than buying a new one? Crying or apologizing might be the best response for a poor child, but such answer will yield a lower "intelligence" score.

EDUCATION WEEK

Volume VI, Number 11 · November 19, 1986

FairTest Reasserts Conflict-of-Interest Charges

By Marleen Nienhuis

WASHINGTON—A public-interest group that monitors the educational-testing industry last week reasserted its charges that the Educational Testing Service is violating federal conflict-of-interest laws because it also operates the only federally financed national clearinghouse on tests. *(See Education Week, Oct. 8, 1986)*.

The group, the National Center for Fair & Open Testing (FairTest), had been invited by Chester E. Finn Jr., assistant secretary for educational research and improvement, to comment on the E.T.S.'s response to charges the group orginally made in September.

In a letter to John G. Weiss, FairTest's director, about a month ago, Mr. Finn agreed that a more thorough examination of the charges was needed and said he would welcome FairTest's reaction to the E.T.S. response before proceeding with the review.

Mr. Weiss again insisted last week that the E.T.S., the nation's largest manufacturer of standardized tests for education, cannot operate the clearinghouse in an unbiased manner because it has a substantial financial interest in the sale of standardized tests. Last year, the E.T.S. earned revenue of $180 million from such sales.

A private, nonprofit corporation, the E.T.S. has operated the Educational Research Information Center on Tests, Measurements, and Evaluation under an Education Department contract since 1970.

In its response in September to FairTest's charges, the E.T.S. denied that it had manipulated information about standardized tests to further its financial interest, as charged by the public-interest group.

Asserting that the E.T.S. had not refuted the conflict-of-interest charges, FairTest officials again have called on Secretary of Education William J. Bennett to launch an "open and independent" investigation into the E.T.S. contract to operate the ERIC clearinghouse.

In addition, the group has specifically asked that Mr. Finn and his press aide, James Bencivenga, be removed from such an inquiry.

Because Mr. Finn is responsible for overseeing the department's conflict-of-interest regulations, Mr. Weiss said, the assistant secretary's participation in a probe of the E.T.S. might result in his "investigating his own behavior." And Mr. Bencivenga should not play a role in the inquiry, Mr. Weiss said, because he had demonstrated "prejudicial attitudes" in the matter. He had publicly called Fairtest "a harassment operation," and accused Mr. Weiss of being "against all testing," Mr. Weiss noted.

Education Department officials could not be reached for comment.

Details of FairTest Charges

FairTest, which has been pressing the conflict-of-interest charges on behalf of some three dozen academics, researchers, and law professors, asserts that the E.T.S. violates federal acquisition regulations that state "an organizational conflict of interest exists when the nature of the work to be performed under a procured government contract may: (a) result in an unfair competitive advantage to the contractor or (b) impair the contractor's objectivity in performing the contract work."

The public-interest group claims that such a conflict exists in principle because the E.T.S. earns more than $180 million a year from the sale of standardized tests at the same time that it controls the information on testing in the only federal databank on tests. Such an arrangement helps the testing service maintain a competitive advantage, FairTest officials maintain. In addition, the group alleges, a conflict in practice is "demonstrated by a clear pattern of bias and self-promotion."

For example, Mr. Weiss said, information supporting the practice of coaching to improve student test scores is either not included or described only minimally in abstracts of library materials provided by the ERIC clearinghouse on tests, measurements, and evaluation. Such a practice is followed, he asserted, because students who are coached skew test scores and undermine the validity of standardized tests.

Publications critical of the sexual and cultural bias in standardized testing are often not included or difficult to find at the ERIC clearinghouse, according to Mr. Weiss.

In addition, he said, most of the informational abstracts on testing refer to other E.T.S. publications. And information on so-called truth-in-testing laws, which allow test-takers to receive scores and questions and answers—and, in some states, to challenge them—is either not available or difficult to obtain through the clearinghouse.

Mr. Weiss noted that the E.T.S. has "spent hundreds of thousands of dollars lobbying against proposed truth-in-testing laws." He added that the E.T.S. also does not make it clear in the clearinghouse database that, when pressured by truth-in-testing legislative initiatives in 26 states, the testing service reversed its opposition to them and agreed to "voluntarily" make testing information available nationwide.

The E.T.S. is currently operating the clearinghouse under a two-year, $530,000 grant from the Education Department. Its president, Gregory R. Anrig, has said the organization provides an annual subsidy of about $70,000 to the testing center.

Since 1981, the company has submitted the only bid for the contract, according to an Education Department spokesman.

The spokesman added that the department tends to select universities and educational institutions whose specialties coincide with those of the 16 ERIC centers.

But the spokesman could not recall another ERIC sponsor who derived substantial income from the sale of a product related to an ERIC subject area.

IN THE NEWS

New commission being formed

A Blue Ribbon Commission is being formed by Bernard R. Gifford, Dean of the Graduate School of Education at the University of California at Berkeley. The Commission will study how standardized tests are used to allocate "educational, training and employment opportunities, as well as many other social and economic benefits" (*Education Week*, December 17, 1986).

Math teaching being harmed

John A. Dossey, President of the National Council of Teachers of Mathematics, argues that America's emphasis on tests has had a negative impact on the teaching of mathematics. "They take more time away from instruction, and they test the wrong things" (*Education Week*, January 21, 1987).

Tests as indicators

Chris Pipho, Education Commission of the States Clearinghouse Director, writes that the time has come to question the use of standardized tests, especially the SAT and ACT, as indicators for policy decision-makers (*Education Week*, January 28, 1987).

Defective state bar exam

Thirty people who were told last November that they had failed the New York State bar exam were subsequently given passing marks, according to the *New York Times* (January 14, 1987). When one test-taker filed an appeal about a question which he claimed had been incorrectly marked, the State Board of Law Examiners decided that there were two correct answers to the question. In addition to the 30 who passed due to the change in scoring, 22 others had their scores raised to the point where they could appeal their grading.

Teacher tests

- The Texas State Legislature has eliminated the three-year-old law requiring teachers to take exams in the subjects that they teach.
- The Georgia Association of Educators, an affiliate of the NEA, has filed a class action lawsuit charging the Georgia teacher certification tests discriminate against blacks and are not job related.
- The Buffalo (NY) Board of Education released statistics in February which showed that 74% of minority candidates who took the district's teacher exam (ETS' PPST) last year failed the test. Only 24% of the white candidates who took the test at the same time failed. The continued use of the test is currently under litigation.

Controversial police exam

After intense criticism and controversy, the Wilmington Police Department (Delaware) has suspended its police promotional exam during the first year of its administration. The exam was published by Landy, Jacobs and Associates of State College, Pennsylvania.

'Exam scam'

A former Metropolitan District Commission Police captain in Massachusetts has been accused of systematically lowering scores of black police officers on promotional exams during eight years of break-ins at civil service offices in Boston, beginning in 1978. This was one charge made in a U.S. District Court hearing into allegations that certain police officers gave or sold exams for $3,000 a copy to colleagues and firemen whose promotions they favored, while downgrading others over a period of several years. With officers of their choice in key positions, they allegedly planned a number of criminal operations, including a $1.5 million bank robbery in Medford.

focus on ETS

- **ETS has filed a $50,000 lawsuit against a California student.** ETS alleges that Bruce Zinky arranged for high school students to take the SAT for other students at fees of $100 each.
- ETS has recently developed a number of new tests, including new competency tests for use in **secondary schools**. It is also working with a consortium of states to set up a State-Sponsored Testing Program to develop **small-volume tests for teacher certification.**
- ETS has brought out **software programs to coach students** for the GRE and GMAT exams. Late last year it introduced a new videodisc--'On your own: preparing to take standardized tests'--which is being marketed to high schools for $150.

International ETS:

- ETS, currently at work in 140 countries, is expanding its international operations. It is helping to design an adult education system and testing center for the **Sultanate of Oman** and is assisting the government of **Honduras** to establish an assessment program for elementary school students. In **Bermuda**, the company is continuing a decade-long involvement by advising on the construction of a new secondary school curriculum. Its Middle East field of operations includes **Jordan**, where it is helping to evaluate the Arabic Reading Series Project. In the Far East, the **Seoul** Olympic Organizing Committee is testing its staff with ETS' Test of English for International Communications (TOEIC). TOEIC is widely used in **Japan, Hong Kong, Indonesia, Korea, and China.**

Occupations that require licensing and certification exams (partial listing)

- Accountants
- Aerobics Instructors
- Arborists
- Architects
- Attorneys
- Auto Mechanics
- Babcock Test Operators
- Barbers
- Blood Bank Specialists
- Boiler Engineers
- Cash Managers
- Chemists
- CPAs
- Chiropractors
- City Planners
- Cosmetologists
- Cytotechnologists
- Dental Hygienists
- Dietitians
- Electrologists
- Engineers
- Firefighters
- Food Service Managers
- Foreign Service Officers
- Golf Pros
- Hairdressers
- Hearing Aid Dealers
- Hemotologists
- Insurance Adjusters
- Insurance Brokers
- Insurance Consultants
- Land Surveyors
- Medical Lab Technicians
- Medical Record Admin.
- Medical Record Tech.
- Microbiologists
- Motion Picture Projectionists
- N.Y. Stock Exchange Brokers
- Nurses
- Nursing Home Administrators
- Occupational Therapists
- Oil Burner Servicemen
- Optometrists
- Osteopathic Physicians
- Pest Control Operators
- Pesticide Applicators
- Physical Therapists
- Physician's Assistants
- Picture Framers
- Plumbers
- Podiatrists
- Police Officers
- Practical Nurses
- Professional Engineers
- Psychological Examiners
- Psychologists
- Radiologic Technologists
- Real Estate Brokers
- Real Estate Salesmen
- School Bus Operators
- School Teachers
- School Principals
- Securities Agents
- Securities Salesmen
- Sewage Treatment Plant Oper.
- Social Workers
- Veterinarians
- Welders

- ETS' **Multistate Cosmetology Licensing Exam** will be given to approximately 25,000 would-be cosmetologists in Hawaii, Illinois, Minnesota, Virginia, and Wisconsin next year. ETS' multiple-choice cosmetology exam has captured 30% of the national market in just the past few years. The exam places special emphasis on quality nail care. ETS has also tailored special tests for states wanting to license "**shampooists and manicurists**."

- The International Dance Exercise Association Foundation (IDEA) contracted with ETS to develop and administer a three-and-a-half hour, 175-item multiple-choice exam for certification of **aerobics instructors**. When a pilot exam was administered at the IDEA convention last summer, one-third of the candidates failed. In October an estimated 3,000 candidates took the test, and IDEA is now examining ways to introduce it into its 33 other member countries.

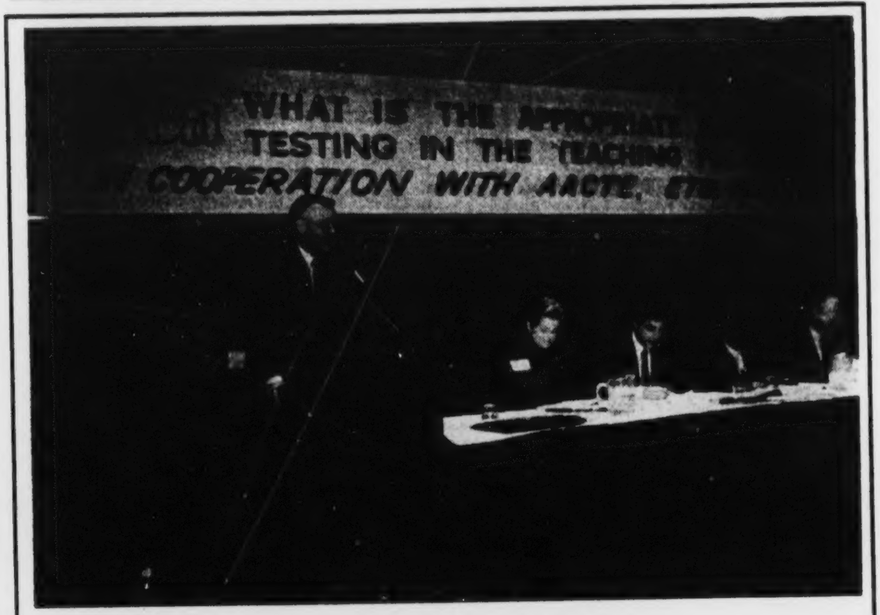
AACTE resolution

In late February 1987, the American Association of Colleges for Teacher Education (AACTE) endorsed the following resolution:

"Be it resolved that AACTE supports the following conditions for the development and use of standardized tests used for admission to or graduation from teacher education programs and for the certification of teachers:

- Test developers should provide researchers and test users with statistical information about their test's reliability and validity. This information shall include an analysis of different ethnic and gender groups' performance on the test as a whole, as well as on each item on the test.
- Truth-in-Testing policies that include the disclosure of test items should be extended to these tests. Such policies should be modeled after those used for other standardized exams such as the Scholastic Aptitude Test and Law School Admissions Test.
- Tests should be constructed so that, insofar as possible, individual or group differences in test scores result from real differences in people's knowledge of the subject being tested and not from inappropriate characteristics of the questions themselves. Whenever possible, and without lowering the validity or reliability of the test for any group, and without altering the difficulty level of the test or its content coverage, tests should be constructed from items that display the least difference in passing rates between minority and majority examinees of comparable backgrounds and educational level.
- A.P.A. *Standards for Educational & Psychological Tests* should be adhered to in the development of these tests."

AACTE represents 717 colleges of teacher education across the nation.



-- AACTE President Norene Daley, FairTest's John Weiss, NEA's Sharon Robinson, and ETS President Greg Anrig.

On December 12 in Washington, DC, FairTest, the Educational Testing Service, the American Association of Colleges for Teacher Education, and the National Education Association participated in an "off-the-record" invitational symposium on "What is the appropriate role of testing in the teaching profession?" The speakers were FairTest's John Weiss, ETS President Greg Anrig, AACTE President Norene Daley, and Sharon Robinson of the NEA. The formal proceedings of the event will be published by the NEA.

- **March 25, 1987**--John Weiss will be the keynote speaker at the Ninth Northern New England Educational Tests, Measurement and Evaluation Conference, to be held at Plymouth State College in Plymouth, New Hampshire. For more information, contact Mike Hansen, (603) 271-3186.
- **April 1, 1987**--Peter Garcia of Pan American University, Peter Roos from the META Project, Ron Vera from the Thomas Rivera Center and FairTest's John Weiss will be participating in two seminars on testing's impact on Hispanic students and workers at the 16th National Association of Bilingual-Bicultural Education Convention in Denver. For more information, contact (202) 822-7870.
- **April 22, 1987**--Robert Linn, a psychometrician at the University of Illinois, Mary Ann McLean, legislative director to New York State Senator Kenneth LaValle, and FairTest's John Weiss will participate in a National Council on Measurement in Education Symposium on "Fairness in Testing: Has the Industry Done Enough?" For more information, contact Professor Kearney, (313) 426-5174.
- A special issue of the *Negro Educational Review*, edited by Professor Asa Hilliard, will be published this spring. FairTest's John Weiss contributed an article on the Golden Rule and Truth-in-Testing reforms. To order, contact Dr. R. Gran Lloyd, Box 2895, General Mail Center, Jacksonville, FL 32203.

Job Openings:

Organizer/Research Associate: Implement advocacy and public education campaigns. Requirements: Commitment to equal opportunity; at least two years experience working on civil rights or educational issues; strong writing/editing/speaking skills; very energetic and organized. Salary/benefits: $16-20,000 per year.

Director: Test-Takers Bill of Rights: Launch FairTest's constituency services program. Mobilize this constituency to promote our advocacy and public education campaigns. Requirements: Commitment to equal opportunity; willingness to work long hours; strong entrepreneurial/marketing/editing/speaking skills; two-year commitment. Salary: competitive.

Office Manager/Research Assistant: Manage office, help with publications, handle conference logistics. Requirements: strong office skills. Salary: depending on experience. Flexible work schedule.

College Internships and Work-Study Positions are also available.

Starting date for all positions: Summer 1987.

To apply: send cover letter highlighting interest, skills and qualifications, a resume and a list of references to FairTest, PO Box 1272, Cambridge, MA 02238.

Experts Available

FairTest maintains a list of experts available to testify in court cases and make presentations before policy-making bodies and civic groups. For more information, contact FairTest.

CAN YOU PASS THIS TRICKY TEST?

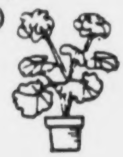
The following standardized test was administered to FairTest Conference participants.

1) What are you supposed to do if you find someone's wallet or pocketbook in a store?
a) Return it
b) Give it to the store owner
c) Keep it; make believe you didn't see it.
--WISC-R IQ Test

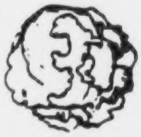
2) How the ________ roses flush up in the cheeks!
a) White b) Pretty c) Small
d) Yellow 3) Red
--Lorge-Thorndike Intelligence Test

3) Which plant needs the least amount of water?

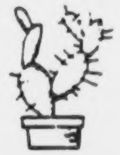
a)



b)



c)



--Stanford Achievement Test

4) In New York City, it is almost impossible to find a place to ________ your car.
a) drive b) repair c) wash
d) park e) hide
--Cognitive Abilities Test

5) Bow:Violin
a) music:piano b) brass:trumpet
c) note:flute d) string:guitar
e) drumstick:drum
--Scholastic Aptitude Test

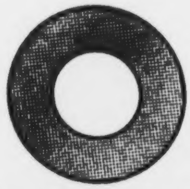
6) Which row contains both the square of an integer and the cube of a different integer?
a) 7 2 5 4 6
b) 3 8 6 9 7
c) 5 4 3 8 2
d) 9 5 7 3 6
e) 5 6 3 7 4
--Scholastic Aptitude Test

7) Most authorities consider intelligence to be influenced by which of the following?
a) Environment alone
b) Primarily environment
c) Heredity alone
d) Primarily heredity
e) Environment & heredity
--ETS' GRE Education Test

8) Which of the following would be the most appropriate intelligence test to use in assessing the intellectual functioning potential of a 6-year-old child who has limited exposure to educational resources and material?
a) Stanford-Binet Intelligence Scale
b) Weschler Intelligence Scale for Children (WISC)
c) Lorge-Thorndike Intelligence Test
d) Otis-Lennon Mental Ability Test
e) California Test of Mental Maturity
--ETS' GRE Education Test

9) What does R.S.V.P. mean?
--Washington, D.C. Police Promotion Exam

10)



In the figure above, what is the greatest number of nonoverlapping regions into which the shaded region can be divided with exactly two straight lines?
a) 6 b) 5 c) 4 d) 3 e) 2
--Scholastic Aptitude Test

STOP

For the "wanted' answers and background material on each test and question, return this test (with the completed subscription form) to FairTest, Box 1272, Harvard Square Station, Cambridge, MA 02238.

Subscribe to the *FairTest Examiner* !

The *FairTest Examiner* will keep you up-to-date on the growing movement for fair & open testing. You will be able to do more than just read about FairTest campaigns--you can help make them successful.

PEANUTS / by Charles M. Schulz

TRUE...FALSE...
TRUE...TRUE... FALSE...TRUE...
MA'AM?

WHAT DO WE DO IF WE COME ACROSS A HALF-TRUTH?

Subscribe today: only $15
(make out checks to FairTest)

Enclosed is my check for $______ to cover the following publications shipped to me postpaid:

____ *FairTest Examiner* ($15/year)

____ Biannual updates on the following topics ($7/each)

- ____ Golden Rule Reforms
- ____ Truth-in-Testing Legislation
- ____ Coaching
- ____ Sex Bias in Testing
- ____ Teacher Testing
- ____ University Admissions Tests
- ____ K-12 Testing

____ 100-page *Source Book* on the Golden Rule Reforms ($20)

____ *The Reign of ETS* by Allan Nairn and Ralph Nader ($30)

NAME______________________________ **TITLE** ____________________

ADDRESS____
(street no.) *(city/town)* *(state)* *(zip)*

Please mail your check and this form to: Box 1272,Harvard Sq. Station, Cambridge, MA 02238

FairTest
National Center for Fair & Open Testing
P.O. Box 1272
Harvard Square Station
Cambridge, MA 02238
(617) 864-4810

Bulk Rate
U.S. Postage
PAID
Cambridge, MA
Permit no. 53278

FairTest
EXAMINER

A Publication of the National Center for Fair & Open Testing — Volume 1, Number 2 — Summer 1987

N.Y. Investigates Test Bias: Panel to Probe Admissions

In late June 1987, the New York Senate and Assembly passed landmark legislation establishing a state commission to examine college and university entrance tests for racial, cultural and gender bias. The plan covers the nation's most popular admissions exams, including the Scholastic Aptitude Test (SAT), American College Testing Program Assessment (ACT), Graduate Record Exam (GRE), Law School Admissions Test (LSAT), Medical College Admissions Test (MCAT), and the Graduate Management Admissions Test (GMAT), among others.

For the first time, test manufacturers will be required to calculate correct-answer rates on each question for whites, blacks, Asians, American Indians and Latinos, as well as for men and women. This "Item Analysis" data will be turned over to the new commission which will examine each question for fairness and recommend how to reform the tests.

A coalition of civil rights, education, feminist, student and consumer groups organized by the New York Public Interest Research Group (NYPIRG) led an aggressive campaign to adopt the proposal. Chief sponsors were Republican-Conservative Sen. Kenneth LaValle, the chair of the Higher Education Committee and sponsor of the state's original Truth-in-Testing law, and Assembly Deputy Speaker Arthur Eve, a leader of the Black and Puerto Rican Caucus.

NYPIRG Lobbyist Blair Horner explains NY testing proposals at FairTest's news conference.

Report Timely

The successful campaign was aided by the release of a joint FairTest/NYPIRG report demonstrating that twice as many boys as girls received the state's prestigious Empire and Regents Scholarships, both of which are awarded solely on the basis of SAT or ACT scores. Initially, testing companies, led by ETS, opposed the bill. But faced with the strong NYPIRG coalition and new data on the exams' discriminatory impact, ETS dropped its opposition to the bill.

The item analysis proposal is part of a broad testing reform package supported by the NYPIRG coalition. Three bills strengthening New York's Truth-in-Testing law passed in 1986. In the closing weeks of the 1987 session and in 1988, the groups will press for extension of Truth-in-Testing to professional licensing tests, a new formula for awarding state scholarships, application of the Golden Rule Principle to exam construction, and controls over the use of experimental sections without test-takers' consent. (See pages 4-5 for a summary of the New York legislative package.)

CONTENTS

Sex Bias in Testing... pp. 2-3
Legislative Reforms.. 4-5
Nation's Report Card. 6
Teacher Testing....... 7
Work for FairTest..... 10
On Golden Rule....... 11
Test Answers.......... 12-13
Order Form............. 15
SAT Flunks Out....... 16

Sex Bias in Testing: Why Women Lose Out

Sex bias in the Scholastic Admissions Test (SAT) and other college admissions tests costs young women millions of dollars in "merit" scholarship aid each year, according to a new FairTest report written by Phyllis Rosser. The FairTest study, *Sex Bias in College Admissions Tests: Why Women Lose Out*, was endorsed by National Organization for Women (NOW) President Eleanor Smeal. It concludes that the failure of these exams to accurately predict females' academic performance denies high school girls fair access to competitive universities and opportunities for special gifted and talented programs.

Extensive Coverage

The FairTest report was released at a New York City seminar in April on *Strategies to Eliminate Sex Bias from Standardized Tests*. Following the meeting, a news conference was held to answer journalists' questions about the report. Press coverage was substantial, with more than 200 different media outlets from 42 states covering the story, including the *New York Times, Christian Science Monitor, Newsweek, ABC Radio, National Public Radio* and *Cable Network News*.

Tests Underpredict Girls' Abilities

Nearly 1.5 million high school students annually take Educational Testing Service's SAT, America's oldest and most widely-used college entrance exam. The gap between the sexes on the SAT is now at its widest margin ever: 61 points in favor of males. Last year girls scored 50 points lower on the Math section, and 11 points lower on the Verbal part of the exam.

What makes this gap particularly worrisome is that ETS' sole scientific justification for the SAT is its supposed ability to predict high school students' first-year college grades. Yet females earn higher grades than males in both high school and college, even during their freshman year. If the SAT predicted equally well for both sexes, girls would score about 20 points higher than boys, not 61 points lower.

High school girls also register lower scores on the two other nationally-administered exams taken by students applying to college--ETS' Preliminary Scholastic Aptitude Test (PSAT) and American College Testing Program's ACT exam. On the PSAT, annually taken by more than one million high school juniors, girls average 53 points (in SAT terms) lower than boys. On the ACT, taken by nearly one million students each year, predominantly in the Mid-West and South, women averaged 6.0 units lower than men in 1986.

Test Scores and Scholarships

More than 750 government and private scholarship agencies rely on the SAT, PSAT, or ACT to evaluate students. These agencies' reliance on test scores results in substantially less scholarship funds for females. In 1985-86, for example, girls received only 36% of the $23 million awarded by the National Merit Scholarship program. Only students who score in the top 1/2 of 1% in their state on the PSAT are eligible to apply for these grants which can be worth up to $8,000 per year. The percentage of female National Merit winners has decreased in recent years, from 40% in 1983-84, to 38% in 1984-85 to the present 36%.

72% Winners Male

New York State annually awards $40 million to high school students who obtain high SAT and ACT scores. Last year, 72% of the winners of the prestigious Empire State Scholarships of Excellence, worth up to $10,000 each, were awarded to males. Following the release of these statistics by FairTest and NYPIRG, the New York Board of Regents sought legislative authorization to change the procedures for awarding scholarships.

Double Jeopardy

"Black, Hispanic and Asian women face double jeopardy because of both the gender and racial biases on these exams," noted FairTest Board member Chuck Stone, who once served as ETS' Director of Minority Affairs. "In every ethnic and racial category males outscore females."

Advocacy Campaign

The FairTest report was strongly endorsed by civil rights, consumer and student leaders. "Biased exams must either be overhauled or discarded," NYPIRG lobbyist Blair Horner told the FairTest seminar. Leslie Wolfe, director of the Project on Equal Education Rights added: "Unless needed changes are instituted, we will have to pursue legislative, legal and advocacy campaigns to guarantee fair access."

As NOW President Eleanor Smeal has stated, "The impact of test bias is devastating to young women. Unfairly low test scores often become self-fulfilling prophecies that destroy girls' self-confidence. The damage can be life-long."

Congressional Hearings on Sex Bias Held

In late April, Congressman Don Edwards (D-CA), chair of the U.S. House Subcommittee on Civil and Constitutional Rights, held hearings on the issue of Gender Bias in Standardized Tests. Witnesses included FairTest consultant Phyllis Rosser and FairTest Board Chair Denise Carty-Bennia, who testified that unfairly low test scores can deter minority women from applying to selective schools or make them feel like second-class citizens when they enroll.

Tests Underpredict GPA

MIT Admissions Director Michael Behnke told the Committee that female applicants to his institution scored significantly lower than male candidates on the Math SAT, Math Achievement Test and on the Science Achievement Test. Such scores would suggest that women would receive lower grade point averages at MIT, he said. But research has found no significant differences in the grade-point averages of women and men at MIT, even when controlling for the courses students took.

"MIT Admissions Director Michael Behnke and Denise Carty-Bennia, Chair of FairTest Board, address Congressional hearings.

Behnke found that because of the lower test scores, MIT's *academic prediction formula* "underpredicts the final GPA for women at all levels."

34 famous men, 1 famous woman

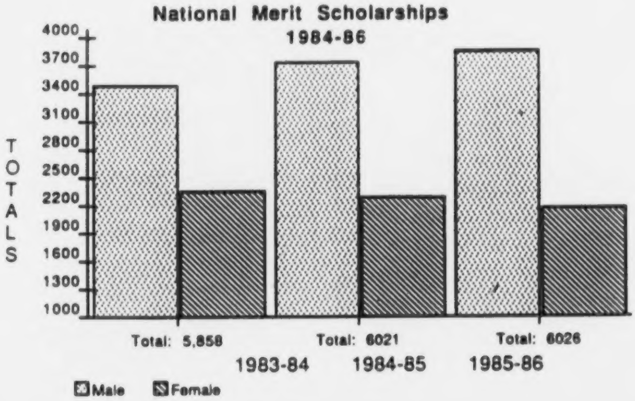
FairTest consultant Phyllis Rosser testified that she was "overwhelmed by the number of males" who populated standardized tests. When females are mentioned, they are usually in such stereotyped occupations as teachers, nurses and secretaries.

Using New York's Truth-in-Testing law, Rosser examined 24 Reading Comprehension passages that appeared on SAT exams in 1984 and 1985. On these exams, thirty-four famous men are mentioned. Eight other men are mentioned, six by name. Only one famous woman, Margaret Mead, is mentioned, and her work is criticized. Two other women are referred to, but not named.

Why Not Change the Test?

Studies by ETS researchers as far back as 1979 show that "females tend to do better on items that have more female or neutral figures than on items in which there are male figures." This means that the male-oriented content is not only offensive; it is also a source of bias.

After listening to the testimony, Rep. Patricia Schroeder, a Colorado Democrat, was moved to ask: "If the tests consistently underpredict females' abilities, why in the world don't they change the test?"



To order:
Sex Bias in College AdmissionsTesting: Why Women Lose Out, 2nd edition, send a check for $7.95 to Fair Test, PO Box 1272, Harvard Square Station, Cambridge, MA 02238.

Testing Reform Goes National

Across the country, more policymakers and citizen groups are launching testing reform campaigns to break down the barriers to access and opportunity posed by unfair exams. Their proposals often apply the FairTest agenda of Truth-in-Testing, Item Analysis, and the Golden Rule Bias Reduction Principle to tests required by the state and local governments or administered within their jurisdiction. The following initiatives are being explored:

In **TEXAS** the Mexican American Legal Defense and Education Fund (MALDEF) has pursued a dual strategy of litigation and legislation to force the state to drop the ETS' Pre-Professional Skills Test (PPST). As the sole criterion for entering teacher training programs, the PPST is essentially the core battery of ETS' National Teacher Exam. Approximately 60% of Mexican American and more than 70% of black applicants failed recent versions of the PPST. Though legislation to bar reliance on this exam did not pass during the legislature's regular session, the state did agree to solicit proposals for a different qualifying exam. MALDEF intends to monitor development of the new test to ensure it is fair, open and educationally sound.

Attorney Don Watkins reports that **ALABAMA** state officials have decided not to appeal a court decision ordering a comprehensive overhaul of the Alabama Initial Teacher Competency Exams constructed by National Evaluation Systems (see *FairTest Examiner* #1 for details of the litigation). Watkins and his colleagues will now turn their attention to legislation to extend Truth-in-Testing and other reforms to additional state-mandated tests which discriminate against minorities.

The **RHODE ISLAND** affiliate of the American Civil Liberties Union has filed legislation establishing a state commission to conduct a detailed Item Analysis of the National Teacher Examination (NTE) which the state uses for pre-classroom certification. The bill, which is also backed by the National Urban Coalition and other reformers, passed the House of Representatives on an overwhelming roll call vote and is expected to clear the Senate with equal ease. The legislative campaign grew out of an excellent Rhode Island ACLU report, *"A Cultural Analysis of the National Teachers Examination and its Impact on Minorities"* issued last fall (copies of the 17-page report are available from FairTest for $2 postpaid).

The **WISCONSIN** legislature is likely to consider three testing reform proposals this fall. Assemblyman Dismas Becker, chair of the Housing and Securities Committee, has submitted a bill to extend the Golden Rule Procedures which now apply to the Illinois Insurance Agent Licensing Test to Wisconsin's version of that exam. A similar bill was approved by committee last year but did not reach the Assembly floor. Also in Wisconsin, Senator Lynn Adelman has drafted two bills to apply basic Truth-in-Testing protections to all admissions and professional licensing exams.

A bill to extend the Golden Rule principle is also pending in **MASSACHUSETTS** where sponsors include Insurance Committee Chair Senator Linda Melconian, Black Caucus Leader Representative Saundra Graham and Senator John Houston, a testing reform activist during his college years. The bill, which is also backed by State Insurance Commissioner Peter Hiam, has attracted the opposition of ETS and insurance industry lobbyists.

SAT Flunks Out (from page 16)

Removing barriers

Middlebury College changed its policy on the SAT to make Middlebury accessible to a wider range of students, especially those from rural backgrounds who "have never heard of the SAT, much less any sort of crash course in how to take the test," College spokesman Herbert Dalron told the *Chronicle of Higher Education*. "Middlebury is going to remove impediments and barriers, rather than erect them." Middlebury will allow students to submit ACT or five Achievement test scores in place of the SAT.

Last year, three other selective institutions examined their need for standardized test scores. As a result of these studies, Bates College no longer requires the SAT, Harvard Business School dropped the GMAT, and Johns Hopkins Medical School dispensed with the MCAT.

100 New York Groups Endorse Legislation

Approximately one hundred organizations and their leaders have endorsed testing reform proposals before the New York State Legislature. Among the key groups in the coalition:

- NYPIRG
- Black/Puerto Rican Caucus
- National Organization for Women
- Parent Teachers Association
- League of Women Voters
- NY Education Association
- NYStudent Association
- Albany State University Black Alliance
- College of Staten Island Shakespearean Society

New York State Testing Reform Legislation

Arthur Eve (l), D-Buffalo, who is the Deputy Majority Leader of the NY Assembly, and Kenneth LaValle (r), R-Port Jefferson, the chair of the NY Senate Higher Education Committee, are leading the fight for testing reform.

Bill #1: Facilitate public oversight of tests by requiring that manufacturers make public all challenges by test takers along with the company's responses (passed in 1986).

Bill #2: Make it easier for students to exercise their Truth-in-Testing rights by requiring test makers to allow students to request scored copies of their answer sheets when they register for the exam (passed in 1986).

Bill #3: Apply Truth-in-Testing to the Test of Standard Written English (TSWE) which test manufacturers had claimed was exempt from public disclosure (passed in 1986).

Bill #4: Require test manufacturers to compile Item Analysis data on college and university admissions tests and submit it to a state commission which will review exams for fairness and make reform recommendations (passed in 1987--see page 1).

Bill #5: Replace the current test score-based system for awarding $40 million in state scholarships with a new formula that better reflects academic ability and performance (pending).

Bill #6: Extend Truth-in-Testing procedures mandating disclosure of exam contents to test-takers and independent researchers to professional and occupational licensing tests (pending).

Bill #7: Apply the Golden Rule Bias Reduction Principle which makes exams fairer, not easier, to new test construction (pending).

Bill #8: Prevent the use of test-takers as unwitting "guinea pigs" by barring experimental sections without their informed consent (pending).

Prop. 48 Penalizes Black Students

According to a NCAA survey, a large proportion of the 650 athletes who failed to meet minimum academic standards laid down for Division I institutions under Proposition 48 are black. Of the 424 athletes who met grade-point average and core curriculum requirements but fell short of necessary SAT and ACT scores, 299 are black. In 1984, a NCAA-commissioned study revealed that had the new requirements been around in 1977, 70% of black athletes who had graduated or were still studying in 1983 would have never been admitted to college in the first place. According to Joseph Johnson, president of Grambling State University, the new survey proved what he and other black-college presidents had been saying for years: "It is a racist rule that was instituted by racist people intent on denying black kids an education."

Test Scores Irrelevant

According to the *Chronicle of Higher Education* for February 11th, a recent ETS report found that selective colleges' graduation rates are not adversely affected by admitting black students with lower-than-average SAT scores. According to the director of black and minority affairs at Swarthmore College, Vernon T. Savage, "This reinforces the notion that the SAT is not a good predictor of academic success."

For copies of testing reform bills, send FairTest $6.00

NAEP to be Transformed

The Nation's Report Card: Improving the Assessment of Student Achievement, Report of the Study Group with a review of the report by a Committee of the National Academy of Education is available from the National Academy of Education, Harvard Graduate School of Education, 108 Longfellow Hall, Cambridge, MA 02138 for $9.00 a copy.

Some twenty years after it was established to see how students nation-wide were performing in reading, writing, math and other subject areas on an occasional basis, the federally-funded National Assessment of Educational Progress (NAEP) is poised for a ten-fold expansion. If a study group appointed by Secretary of Education Bennett gets its way, NAEP will grow from a $4 million to a $26 million a year enterprise, testing 700,000 students at a throw instead of the current 70,000. The new, "improved" NAEP would assess a whole new range of subject areas and types of skills, and provide data which can be used for state-by-state, school district-by-school district and even school-by-school comparisons.

In short, the revamped NAEP would be a huge new report card for comparing K-12 educational systems. This is a total turnaround from NAEP's original intent.

When NAEP was established 20 years ago, state-by-state comparisons were rejected out of hand as an encroachment on state and district educational autonomy. Furthermore, officials feared that such comparisons would lead school districts to tailor their curricula to boost NAEP scores.

"A runaway freight train"

Why is this expansion into territory once regarded as firmly off limits suddenly on the agenda? In its March report, *The Nation's Report Card*, the NAEP Study Group refers somewhat vaguely to "rising public demand." Twenty years ago no one wanted to compare the states but now "just about everyone agrees that there ought to be" such comparative data, and that NAEP is the body to supply it. No mention is made of the continued opposition to such proposals expressed by bodies like the Joint Directors of Research and Evaluation and National Association of Test Directors Committee, which observed in a 1986 report that "the recent movement toward state-by-state comparisons may be akin to a runaway freight train...which would ultimately lead to the compilation and distribution of extremely questionable data."

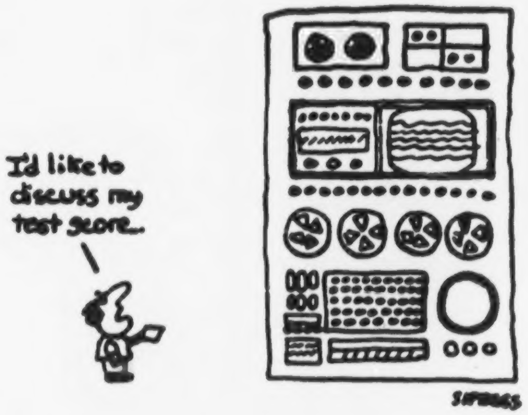
Are we in fact ready for the "first-class report card" promised us by the Study Group? And is NAEP, currently under contract to ETS, really up to the job?

The Need to Go Slow

In its review accompanying the *Nation's Report Card*, the National Academy of Education urges "careful deliberation," citing possible negative consequences in the schools, the danger that NAEP will sacrifice quality if it spreads its efforts too widely, the "real danger of getting lost in the numbers" and of the misinterpretation of data and overzealous use of test results. Linda Darling-Hammond of the Rand Corporation, a member of the Study Group, added this critical warning:

> *There is a delicate balance between developing first-rate assessment of what the nation's students know and can do and attempting to negotiate a multi-purpose testing and data collection effort that may satisfy many objectives superficially but none of them very well.*

What is needed now, concludes the National Academy of Education, is further reflection so that NAEP could play a role in improving education. The Academy observed that "we have already witnessed many school problems precipitated by testing programs that were poorly conceived and implemented" and expressed the hope that NAEP could help states and districts avoid "such pitfalls" in the future. The unstated implication is clear: if NAEP does not take care to become part of the solution, it might turn into just another problem.


I'd like to discuss my test score..
SIPRESS

Teacher Assessment Assessed

A Stanford University research team, backed by a grant of $817,000 from the Carnegie Corporation, is looking into innovative methods of evaluating teachers for the new national teacher certification board proposed by the Carnegie Forum on Education and the Economy. The team believes that the new tests could "sound the death knell for many tests now in use."

According to the project's associate director Gary Sykes, who is professor of teacher education at Michigan State University, existing tests have "tended to trivialize and reduce teaching to a few measurable but 'scientifically validated' kinds of skills...everybody is recognizing that a lot of that just sells teaching terribly short." The project's director, Lee S. Shulman of Stanford University, claims that emphasizing a teacher's performance in the classroom--and not on a multiple-choice test--should help reduce minority failure rates. In Shulman's view, the new assessment procedures, envisaged as a "library of models" which will continually change, should measure such qualities as "wisdom and practice, the release of enthusiasm and energy."

California and Connecticut are working together with the Stanford University team jointly to develop new teacher-licensure tests that go beyond the pencil-and-paper approach to measure what beginning teachers know in their subject area.

New ETS-Created Teacher Council

ETS' Teacher Programs Council recently held their first meeting under the chairmanship of Michael Timpane, president of Teachers College of Columbia University. The 20 members of the Council were selected by ETS. Council creation has a long history at ETS. A former ETS official explained it as an important part of ETS strategy:

> *"Bob Solomon...had a brilliant managerial stroke in 1966. ...He had the foresight to see that ETS might be coming under the gun for different accountability contingencies and cries for closer scrutiny in the future. So Bob went ahead and initiated the idea of setting up the various councils and boards with very elaborate internal checks and balances which would hinder them from coming out and challenging ETS. Then, if someone said to ETS, say, you monolithic testing Goliath you, who do you answer to?' he could point at the boards and councils and say, 'We answer to them'"*
>
> --A. Nairn, *Reign of ETS*, p. 304.

St. Louis Teacher Assessment Under Attack

A new study has criticized the St. Louis' Board of Education's policy of using students' test scores in the formal evaluation of teachers' performance. The report found that the exam, the California Achievement Test, is poorly matched with the district's curriculum. According to the study, which was commissioned by the local teachers' union, on average only 9.5% of all 'content items' were common to the test, the curriculum and assigned text books in the four grades under scrutiny. Twenty-two teachers in St. Louis are currently on probation in large part because of low student test scores. The teachers union, an AFT affiliate, is seeking to have the evaluation policy overturned in court.

A Look at the NTE

"My own content analysis of a sample test for one of the most frequently used tests of professional knowledge (the National Teacher Examination's Professional Knowledge Test) revealed that less than 10 percent of over one hundred questions required knowledge of theory, research, or facts pertaining to teaching and learning...The remaining questions required knowledge about testing and assessment (10 percent), knowledge about school law and administrative procedures (25 percent), careful reading or knowledge of simple word definitions (15 percent), agreement with the test's teaching philosophy (25 percent) or agreement with the test's definition of socially or bureaucratically acceptable behavior (15 percent). About 40 percent of the questions, by my reckoning, did not have a "right answer"...Most of the knowledge and judgement involved in good teaching is not easily boiled down to a multiple-choice question with one simple answer...all of the tests I have seen are limited in their measurement by the scarcity of important teaching questions answerable in multiple-choice formats; the questions with clear, correct answers are not very important or profound."

Linda Darling-Hammond, "Teaching Knowledge: How Do We Test It?" *American Educator*, Fall 1986.

No New Teacher Tests

NEA Today, the official publication of the National Education Association, claims that the push by states to adopt competency tests for teachers is grinding to a halt. "Tests have turned out to be wasteful exercises that demoralize teachers while adding little, if anything, to the quality of education."

Focus on ETS

ETS as Office Park Developer

ETS, which claims that applying the Golden Rule procedure or reducing test fees is economically unfeasible, is looking for ways to spend its $54 million surplus. The latest plan is to double the office space on its 370-acre Lawrenceville, New Jersey "campus" and become a commercial landlord. In the tight Princeton real estate market, the new office space will rent for more than $10 million per year. According to the *New York Times*, area residents are outraged by the potential impact of ETS' development plans on the local infrastructure and by the abuse of the company's non-profit status.

THE HENRY CHAUNCEY CONFERENCE CENTER
ETS

Office Park of the Future?

ETS Transitions

Terrel H. Bell, former U.S. Secretary of Education, has joined ETS' Board. When he was Reagan's first Secretary of Education, Bell instituted the "wall charts" which ranked states by their SAT and ACT scores... Former Congresswoman **Yvonne Braithwaite Burke** has also joined the ETS Board...In July 1987 **Mari-Luci Jaramillo** will become the director of ETS' Berkeley office. She is currently vice-president for Student Affairs at the University of New Mexico. Dr. Jaramillo is president of the Board of the Thomas Rivera Center, as well as a member of the boards of the College Board, the Children's Television Workshop, and the National Council of La Raza...**William Turnbull**, president of ETS from 1970-81, died on May 21 at the age of 67.

New Title 1 Classification System Developed

ETS' Jerry Jenkins has developed a system which classifies and rank orders children according to basic skills deprivation. The system was created in response to a request from a Florida school district which was having problems accurately placing children into Chapter 1 programs. According to Jenkins, most school districts use single measures, such as the latest standardized test score, for classifying and selecting individual children. In many cases, such placements are not very accurate.

Testing for Jobs

In response to a letter from the American Psychological Association, Secretary of Labor William E. Brock has stated that scores on ETS' General Aptitude Test Battery should not be used to deny people with disabilities access to jobs. Currently the GATB is being used for all 12,000 jobs listed in the *Dictionary of Occupational Titles* in all but five states.

International Testing

ETS' Test of English for International Communication will have more than 100,000 examinees in **Japan** this year and is expanding at a rate of more than 15% per year. ETS' Test of English as a Foreign Language (TOEFL) program is designing a new system of computerized registration for use in the **People's Republic of China**, where 20,000 students took the test last year. Plans are in the works to introduce similar systems in **Taiwan**, **Japan** and **South Korea**.

In **South Korea** officials have "changed university entrance requirements to de-emphasize standardized tests, which encourage rote memorization, and rely more on essays, interviews, and high school grades," according to the *Chronicle of Higher Education*.

During the general election campaign in **Britain**, Prime Minister Margaret Thatcher told reporters that under the Tories, state schools would be encouraged to impose tuition fees and select their students through a series of tests.

Excellence Test

Senator Claiborne Pell of Rhode Island, chair of the Subcommittee on Education, Arts and the Humanities, introduced in early April Bill S.913 which would require the Department of Education to create a "comprehensive test of academic excellence" for high school seniors in order to "identify students with outstanding aptitude and ability." Senator Pell's initiative does not define "aptitude" or "ability," or describe the test's format or content.

Coaching

Testing for Tests

A testing and a coaching company are both bringing out exams intended to help students prepare for college admissions tests. The American College Testing Program has unveiled the P-ACT+, a preview of its ACT exam aimed at 10th graders. The Princeton Review coaching company is marketing the Princeton Review Assessment (PRA), designed to predict both SAT and ACT scores and aimed at the 1.1 million juniors who annually take ETS' PSAT. For more information, contact: The P-ACT+ Test Administration, PO Box 168, Iowa City, Iowa 52243; The Princeton Review, 10 Columbus Circle, Suite 1260, NY, NY 10019.

Coaching, ETS Style

ETS, which still claims coaching has no appreciable effect on test scores, has just brought out a videodisk, price $150 per copy, called "On Your Own: Preparing for a Standardized Test." The videodisk gives such valuable test-taking tips as "Get a good night's sleep" and "Wear comfortable clothing."

ETS Launches Software

"Practicing to take the GRE," a new computer software package, has been developed by ETS and is on sale for $55.

Score with Caffeine

Northwestern University researchers looking into the effects of caffeine on complex thinking processes gave their subjects two or three cups of coffee and then asked them to answer practice questions from the Graduate Record Exam. They concluded that caffeine was a test-taking aid for "impulsive people," but inhibited the reasoning process in the introverted.

Coaching Camp

For $1,375, 13-18 year olds can spend 10 days in the summer at Super Camp on the Vassar campus and be drilled in how to get better SAT scores.

Worth Reading

"Legitimate Ways to Prepare Students for Standardized Tests," *1986 Symposia* of the National Association of Test Directors, available from NATD, Oklahoma City Public Schools, 900 N. Klein Street, Oklahoma City, OK 73106.

The LSAT Exposed: Tricks from 12 LSAT's and *The GMAT Exposed: Tricks from 20 GMATs*, by David White, available at $10 per copy from Testing for the Public, 1308 Peralta, Berkeley, CA 94702.

Source of Bias

"Whereas Caucasian students performed similarly in familiar examiner and unfamiliar examiner conditions, black and Hispanic children scored significantly and dramatically higher with familiar examiners...At least one facet of test procedure, examiner unfamiliarity, appears to depress selectively minority examinees' performance."

--Douglas Fuchs, *"You Can Take a Test Out of the Situation, But You Can't Always Take the Situation Out of a Test: Bias in Minority Assessment."* (For reprints, write to Douglas Fuchs, Box 328, Department of Special Education, Peabody College, Vanderbilt University, Nashville, TN 37203.)

Police "Exam Scam" Judgement

A Massachusetts jury has convicted six of the seven defendants in the Metropolitan District Police scandal involving the theft of police and fire promotional exams. The exams were sold to officers in nine different police and fire departments for as much as $3,000 a copy. One defendant collected at least $50,000 for selling stolen exams over a six-year period. As well as favoring their friends, the defendants systematically lowered the scores of black police officers. They were sentenced to terms from 5-20 years. Individuals who were unfairly denied promotions are considering filing lawsuits to recover lost wages.

Settlement Clears Way for Police Promotions

After the New York State Supreme Court ruled that at least some of the "wrong" answers on a police promotional exam were as acceptable as those originally deemed "correct", the way was cleared in early February, 1987 for the promotions to the rank of sergeant of 765 police officers in New York City. Shortly after the exam was administered in 1983 civil rights groups charged that it had "disparate impact" since only 1.6% of black officers and 4.4% of Hispanic officers passed the test, compared to 10.5% of white officers. In 1985, the City agreed to promote officers proportional to the number of white, black and Hispanic candidates who took the exam. There were further delays because of lawsuits challenging individual answers to test questions. A group of white police officers denied promotion under the terms of settlement are appealing the ruling.

Events

- On July 9th Yale University will sponsor an evening symposium on Admissions Testing featuring University of Pittsburgh Professor Lloyd Bond and FairTest's John Weiss.
- The 20th Annual Convention of the Association of Black Psychologists is meeting in Atlanta, Georgia on August 18-23. For more information, contact Dr. Charlyn Harper, PO Box 310797, Atlanta, Georgia 30331.
- The Third Biennial Conference on Minority Assessment will meet on Nov. 18-20th, concurrently with the 18th Annual Conference of the Rocky Mountain Educational Research Association in Tucson, Arizona. Contact Karen Gendron, Conference Coordinator, Education Building, Box 604, University of Arizona, Tucson, Arizona 85721.
- The American Educational Research Association and National Council for Measurement in Education will hold their 1988 annual meeting on April 5-9 in New Orleans, LA. Address inquiries and send proposals for papers to the AERA Meeting Coordinator, 1230 17th St. NW, Washington, DC 20036.

Special Issue on Golden Rule

The Summer 1987 *Educational Measurement: Issues and Practices* has a special focus on the Golden Rule Bias Reduction Procedure. Featuring articles by Golden Rule Chief Executive Officer **J. Patrick Rooney**, ETS Researcher **Jane Faggen**, U. of Illinois Professor **Robert Linn**, U. of Pittsburgh Professor **Lloyd Bond** and FairTest's **John Weiss.** It is available for $5.00 from NCME, 1230 17th Street, N.W. Washington, DC 20036.

Work with FairTest

MANAGING DIRECTOR

Duties: Responsible for day-to-day implementation of all FairTest projects, including public education, advocacy, publications and fundraising. Work with Executive Director to design strategy and allocate resources.

Requirements: Strong managerial, planning, fundraising and writing skills; experience working on civil rights and/or education issues a plus.

Salary: competitive. Benefits.

ADMINISTRATIVE ASST.

Duties: Office support, including phones, mailings, correspondence, assisting with publications and logistics for events.

Requirements: strong office, writing and typing/word processing skills.

Salary: $14,000 to $17,000/yr. Benefits. Three weeks vacation.

• • •

INTERNSHIPS and WORK-STUDY positions are also available.

To apply, send us a cover letter and resume as soon as possi-

FairTest Board of Directors

Denise Carty-Bennia, Chair
Professor of Law at Northeastern University
Pamela George
Professor at North Carolina Central University
Paul Pottinger
President of the New Manhattan Corporation
William Robinson
Executive Director, Lawyers Committee for Civil Rights Under Law
Steve Solomon
Attorney, Proskauer, Rose, Goets & Mendelsohn
Chuck Stone
Senior Editor of *Philadelphia Daily News*, and Professor at University of Delaware

FairTest Staff

Amy Allina, Intern
Nancy Murray, Coordinator of Publications and Research
Bob Schaeffer, Public Policy Communications Consultant
Sarah Stockwell, Organizer, Sex Bias Coordinator
John Weiss, Executive Director

FairTest Examiner

The *FairTest Examiner* is a publication of the National Center for Fair & Open Testing (FairTest), P.O. Box 1272, Harvard Square Station, Cambridge, MA 02238, phone (617) 864-4810. We are a national research and advocacy organization working to ensure that the 40-plus million standardized multiple-choice exams administered annually to America's students and job applicants are fair, open and educationally sound.

We would like to thank the Rockefeller Family Fund and the Ford Foundation for suppporting our work.

The APA Monitor
American Psychological Association

'Golden rule' a plus for tests

Dear Editor:

For those who have closely followed the standardized-testing reform debate over the past decade, Educational Testing Service President Gregory Anrig's flip-flop reneging on the Golden Rule legal settlement, which he signed two years ago, ("Golden Rule: second thoughts," January), should come as no surprise.

The misstatements, exaggerations and "red herring" arguments are similar to those ETS and other test manufacturers have used in the past to block consumer protection laws and all other forms of public control over their billion-dollar-a-year business.

In fact, the Golden Rule principle is based on a simple statistical technique that makes tests *fairer, not easier.*

The National Center for Fair and Open Testing (FairTest) and other advocates of Golden Rule do not, as Anrig erroneously states, believe "that group differences in performance on test questions primarily are caused by 'bias.'" Rather, we recognize that test-score differences reflect a host of causes, including differences in test-takers' abilities.

The Golden Rule principle simply enables test publishers to ensure that their exams measure real knowledge differences between test-takers and not irrelevant information.

Under Golden Rule, the same content areas will be covered as on previous tests and the exam will be of the same overall level of difficulty. The only difference is that within groups of equally valid items in the same content areas, test publishers must select those items that display the least difference in correct-answer rates between majority and minority groups.

This point is crucial. Nothing in Golden Rule changes the content specification of future exams. Nothing alters the ability of test administrators to set cut scores or passing rates. If group performances differ widely on all the items in a given content area, these questions will still be used on exams assembled under Golden Rule.

Nor do legislative and judicial remedies advocated by groups like the Mexican American Legal Defense and Education Fund (MALDEF), the Alabama affiliate of the National Educhtion Association (NEA), the Rhode Island Chapter of the American Civil Liberties Union (ACLU) and other civil rights and educational reform organizations seek to blindly apply all the details of the Illinois agreement. Rather than mandating complete adoption of the settlement that ETS President Anrig once accepted, the most recent extension proposals simply seek applications of the Golden Rule principle: that objective statistical techniques must be used to weed out items that exaggerate group differences.

Indeed, application of the Golden Rule settlement in Illinois had just this impact. According to data revealed to the advisory committee established to monitor the agreement's implementation, the gap between the average scores of black and white test-takers closed by 25 percent under the first Golden Rule assembled forms. Since many minorities previously scored just below the passing level "cut-score," the percentage of blacks who become insurance agents should increase significantly. Applying the procedure to future exams should further reduce group differences caused by irrelevant factors.

Finally, Anrig is wrong in stating that FairTest's main effort has been "to advocate state legislation that would mandate the 'Golden Rule' procedure for all educational and occupational tests."

Our organization is committed to the extension of the Golden Rule principle. But we are also concerned with a wide range of reforms designed to ensure that the 40 million standardized, multiple-choice tests administered each year are fair, open and educationally sound. Current FairTest projects focus on eliminating gender bias in ETS's Scholastic Aptitude Test and other exams, extending "Truth-in-Testing" protections to professional licensing tests and blocking the trend toward test-driven instruction in the nation's schools.

Standardized multiple-choice exams have become our nation's cradle-to-grave arbiter of social mobility, used for everything from kindergarten placement to professional promotion.

These exams are far too important to be left to the sole control of those who profit from their sales. Test takers, test users, policy makers and other "consumers" also have a profound interest in the quality of these tests. These groups are FairTest's constituency.

Responsible test manufacturers and psychometricians have nothing to fear from our proposals. We urge those who desire more information about the Golden Rule reforms to contact FairTest at P.O. Box 1272, Harvard Square, Cambridge, MA 02239.

John Weiss
Executive Director
National Center for Fair & Open Testing, Inc. (FairTest)
Cambridge, Mass.

Vol. 18, No. 4
April 1987

The Answers

Here are the 'wanted' answers to the test included in our last issue.

The following standardized test was administered to FairTest Conference participants.

1) What are you supposed to do if you find someone's wallet or pocketbook in a store?

--WISC-R IQ Test

2 point response: Return it
1 point response: Give it to the store owner
0 point response: Keep it. Make believe you didn't see it.

Is a poor child who keeps the money really less intelligent? Harold Dent, chair of the Testing Committee of the Association of Black Psychologists, argues that many minority children might not pick up the wallet or pocketbook for fear of being accused of stealing it. Such a response would earn the child no points.

2) How the ________ roses flush up in the cheeks!

a) White b) Pretty c) Small
d) Yellow e) Red

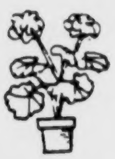
--Lorge-Thorndike Intelligence Test

The "wanted" answer is red, but only if the cheek in question is white.

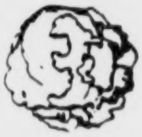
3) Which plant needs the least amount of water?

--Stanford Achievement Test

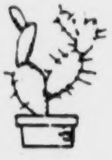
a)



b)



c)



The "wanted" answer is the cactus. Researchers Walt Haney and Laurie Scott, in their study *Talking with Children about Tests*, found that some children selected the cabbage because it had already been picked, and therefore didn't need any water!

4) In New York City, it is almost impossible to find a place to ________ your car.

a) drive b) repair c) wash d) park e) hide

--Cognitive Abilities Test

We've just seen the test, and not the answer key. Your guess is as good as ours.

5) Bow:Violin

a) music:piano b) brass:trumpet c) note:flute
d) string:guitar e) drumstick:drum

--Scholastic Aptitude Test

This question is just one of the numerous SAT items that require students to be familiar with the activities and/or vocabulary of America's upper middle class. The "wanted" answer is (e).

6) Which row contains both the square of an integer and the cube of a different integer?

a) 7 2 5 4 6
b) 3 8 6 9 7
c) 5 4 3 8 2
d) 9 5 7 3 6
e) 5 6 3 7 4

--Scholastic Aptitude Test

ETS initially claimed that the correct answer was (b), since 9 is the square of 3 and 8 is the cube of 2. Answer (c) was included as a distractor, designed to fool students who recognized that 4 was the square of 2 and that 8 was the cube of 2, but failed to remember that the question specified that the integers be different. Then high school student Michael Galligan, using New York's Truth-in-Testing law, challenged the item. He argued that the question had "two correct answers" because 4 is the square of 2 and -2, so both (b) and (c) were correct. ETS had to raise the scores of more than 20,000 students because of the discovery of this error.

7) Most authorities consider intelligence to be influenced by which of the following?

a) Environment alone
b) Primarily environment
c) Heredity alone
d) Primarily heredity
e) Environment & heredity

--ETS' GRE Education Test

Who is "most authorities"? What is "intelligence"? The answer to these questions will determine what answer one marks as correct. "Authorities" have disagreed about the definition of intelligence since Binet created the first intelligence test a century ago. What intelligence is ETS talking about? Is ETS equating intelligence with IQ test scores? Don't think too long about this question; you have only 51 seconds to answer each item. The "wanted" answer is (e).

8) Which of the following would be the most appropriate intelligence test to use in assessing the intellec-tual functioning potential of a 6-year-old child who has limited exposure to educational resources and material?

a) Stanford-Binet Intelligence Scale
b) Weschler Intelligence Scale for Children (WISC)
c) Lorge-Thorndike Intelligence Test
d) Otis-Lennon Mental Ability Test
e) California Test of Mental Maturity

--ETS' GRE Education Test

The "wanted" answer is (b), the Wechsler Intelligence Scale for Children (WISC). The WISC is a three-decades-old exam that many experts believe is biased (see item #1 above). Last year Federal Judge Robert Peckham ruled that the WISC test, as well as a newer version of the exam called the WISC-R, were unfairly biased against black children, and therefore could not be used to track black children into classes for the "educably mentally retarded." Judge Peckham's ruling has been affirmed by the U.S. Court of Appeals, and is currently the law throughout California.

Testing Children Taught at Home

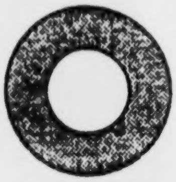
The June 1987 issue of *Growing Without Schools* has a special focus on the testing of children being taught at home. Price $3.50 from Holt Associates, 729 Boylston Street, Boston, MA 02116.

9) What does R.S.V.P. mean?

--Washington, D.C. Police Promotion Exam

Due to a lawsuit brought by the Lawyers' Committee for Civil Rights Under Law, officials acknowledged that this item was not job-related, and took it off the police exam.

10)

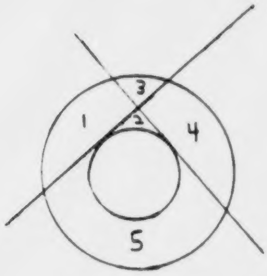


In the figure above, what is the greatest number of nonoverlapping regions into which the shaded region can be divided with exactly two straight lines?

a) 6 b) 5 c) 4 d) 3 e) 2

--Scholastic Aptitude Test

Graduates of coaching courses such as the Princeton Review know that on the SAT easy questions are at the beginning of a section, and difficult questions are at the end. On the actual SAT, as well as on our test, this item was the last question. Thus it is a question designed to elicit wrong answers from most test-takers. Thus, the obvious choice (c) 4 has to be wrong. But if you can can create at least 4 regions, then choices (d) 3 and (e) 2 must also be incorrect. This leaves choices (a) and (b). Since low-scoring students often select the largest number whenever they are given a question asking for the greatest number, well-coached students know that ETS rarely makes the largest number the "wanted" answer to difficult items. Thus, the savvy test-taker selects choice (b), which just happens to be the "wanted" answer (see drawing below).



Boston Sunday Globe

Founded 1872

WILLIAM O. TAYLOR, *Chairman of the Board and Publisher*
JOHN P. GIUGGIO, *Vice Chairman* JOHN S. DRISCOLL, *Executive Editor*
RICHARD C. OCKERBLOOM, *President* MARTIN F. NOLAN, *Editor, Editorial Page*

DAVID STANGER, *Senior VP & General Manager* THOMAS F. MULVOY JR., *Managing Editor*
ARTHUR KINGSBURY, *VP & Treasurer* ALFRED S. LARKIN JR., *Deputy Managing Editor*
MILLARD G. OWEN, *VP Marketing & Sales* HELEN W. DONOVAN, *Deputy Managing Editor*
FRANK E. GRUNDSTROM JR., *VP Human Resources* S. J. MICCICHE, *Assistant Executive Editor*
ROBERT L. HEALY, H. D. S. GREENWAY, *Associate Editors*

Publishers
CHARLES H. TAYLOR, *1873-1922* WILLIAM O. TAYLOR, *1922-1955* WM. DAVIS TAYLOR, *1955-1977*
President JOHN I. TAYLOR, *1963-1975*
Editor L. L. WINSHIP, *1955-1965*
Editor THOMAS WINSHIP, *1965-1984*

Equality in testing

Next Saturday thousands of New England high school students will troop into classrooms, auditoriums and gymnasiums to take a series of tests that are likely to discriminate against half the students taking them.

Because of their bias, these Scholastic Aptitude Tests will have such unfortunate effects on the lives of many of the students that their use should be sharply curtailed.

Over the years, criticisms have been levied at the SATs – that they discriminated against students from lower socio-economic households in general and against minority students in particular.

Now, the National Center for Fair and Open Testing – "FairTest" – has compiled evidence that suggests that the SATs also discriminate against young women, including those from upper-income homes and those who are white.

The facts are not in dispute. Last year, men scored on the average 50 points higher on the math SAT and 11 points higher on the verbal SAT. This disparity has been gradually increasing over the past dozen years, particularly in the verbal section.

The gap, says FairTest, "is surprising, since the justification for the use of the SAT is that it is supposed to predict high school students' first-year college grades. Yet, girls have higher average grades than boys in both high school and college, even during their freshman year."

SAT officials have attempted to explain the disparity away by arguing that women tend to take easier courses in college – fine arts and literature instead of mathematics and science – and thus get better grades even though their SAT scores were lower. That argument has been rebutted by admissions officers at Massachusetts Institute of Technology, who have found that women hold their own across all subject areas, even math and science, despite their lower SAT scores.

Recognizing that the SAT is not a perfect test, most competitive colleges use it merely as one factor – along with high school grades, teachers' recommendations, and even personal interviews – in determining whether to admit a student.

What is troubling, however, is that the prestigious National Merit Scholarships – worth a total of $23 million a year – are awarded solely on the basis of scores obtained on the SATs (usually taken during the senior year) and the PSATs (taken during the junior year). Many other scholarships are awarded to college-bound students largely because they have received a National Merit Scholarship.

FairTest has discovered that because of the apparent sex bias in the SATs, young men are winning a disproportionate number of merit scholarships. Last year, men received 64 percent of the scholarships; in Massachusetts, they received 61 percent.

To counter this finding, SAT officials are arguing that more women, especially those from lower socio-economic groups, are taking the tests, thus bringing down the scores. To argue this, however, is to admit the more basic charges that the SATs are biased against lower socio-economic groups in general.

The weight of the evidence appears to have finally tilted strongly against the SATs. An increasing number of colleges are realizing that, and are finding other ways to judge whether to admit an applicant. The SATS also should not be allowed to remain as a major determinant for scholarships.

Subscribe to the *FairTest Examiner*

The *FairTest Examiner* will keep you up-to-date on the growing movement for fair & open testing. You will be able to do more than just read about FairTest campaigns--you can help make them successful!

Subscribe today: only $15

Enclosed is my check to FairTest for $_____ to cover
the following publications shipped to me postpaid.

____ ***FairTest Examiner***, published quarterly...$15/year
____ ***Golden Rule Bias Reduction Reforms Sourcebook***, (250 pages)...$20
____ ***Sex Bias in College Admissions Tests: Why Women Lose Out***, by Phyllis Rosser and FairTest staff, 2nd edition (40 pages)...$7.95
____ ***The Reign of ETS***, by Allan Nairn and Ralph Nader (550 pages)...$30

FairTest Updates issued bi-annually: Spring & Fall.
(first issues mailed in May 1987)...$7/year per topic:

- ____ *Sex Bias in Testing*
- ____ *Golden Rule Reform*
- ____ *Test Coaching*
- ____ *Truth-in-Testing Legislation*
- ____ *Teacher Testing*
- ____ *University Admissions Tests*
- ____ *K-12th Grade Testing*

____ **TOTAL PAID**

Name ______________________ Title ______________________

Organization ______________________

Address ______________________

______________________ Zip ______________________

Please make checks payable to FairTest and mail to:
FairTest, P.O. Box 1272, Harvard Square Station, Cambridge, MA 02238.

SAT Flunks Out

"A Rejection Slip for the SAT" was the *Newsweek* magazine headline in May following Middlebury and Union College decisions to no longer require students to submit SAT scores, the Brown University student referendum on the SAT and the FairTest report on *Sex Bias in College Admissions Tests: Why Women Lose Out* (see pp. 2-3).

Own Investigations

Unlike most institutions, which rely on College Board or ETS generated studies to determine their need for the SAT, both Middlebury and Union independently investigated their use of the exam. Union College set up a committee of faculty and administrators to look into the SAT after a female applicant complained that she resented "having to spend money that I do not have, and time that I don't have, to take coaching courses to generate a set of scores that I present to you that don't represent me." The committee recommended that Union stop requiring the SAT, and allow students to submit ACT or Achievement Test scores instead.

"Women, minorities and people in the arts--they're people who are very often hurt by SAT scores," concluded Union College President John S. Morris. "We are convinced the SAT cannot escape cultural bias. The evidence we have reviewed is quite clear that the SAT handicaps certain groups, women and minority students being the prime examples. We believe making the SAT optional will encourage applications from students in those groups."

"We have become dismayed at the emotional energy being committed to the SAT by secondary school students and their parents," President Morris continued. "The widespread existence of coaching courses is an example; so, too, is the anguish expressed by bright and serious students who perform poorly on the SAT even though they distinguish themselves in a variety of ways."

Union College is a selective liberal arts college with an engineering emphasis. Last year the College attracted more than 3,000 applicants for the 500 openings in its freshman class. Average SAT scores for recent classes have been in the top 10% nationally.

(turn to page 4)

Michael Spalter (left) and Mark Safire are founders of Students Against Testing, which sponsored the Brown University referendum on making the SAT optional. The proposal fell just 14 votes short out of more than 1,500 cast, and organizers plan to intensify their campaign next year.

FairTest
National Center for Fair & Open Testing
P.O. Box 1272
Harvard Square Station
Cambridge, MA 02238
(617) 864-4810

Bulk Rate
U.S. Postage
PAID
Cambridge, MA
Permit No. 53278

Address Correction Requested

Complimentary Copy.
Please subscribe today.
See page 15.

FairTest
EXAMINER
The National Center for Fair & Open Testing Volume 1, Number 3 Fall 1987

States to Protect Test Takers, Reduce Race and Gender Bias

For the first time in the history of standardized testing, state governments are launching investigations to determine whether widely-used exams unfairly discriminate against minorities and women.

This summer, both Rhode Island and New York adopted laws establishing official study commissions to examine test items for fairness and recommend necessary reforms. New York legislators also overhauled the state's Regents scholarship program, ending the use of SAT and ACT scores as the sole means to determine winners of $40 million in annual awards.

ACLU Examines NTE

The new Rhode Island commission will investigate charges that ETS' National Teachers Examination (NTE), used to certify all classroom educators in the state, unfairly discriminates against minorities. A fall 1986 report, *A Cultural Analysis of the National Teachers Examination and Its Impact on Minorities*, by the state's American Civil Liberties Union affiliate documented the "devastating" effect of the NTE on black and Hispanic applicants and concluded: "There is no compelling reason for using a test that systematically excludes certain classes of people from the teaching profession where there is no evidence that such people make for less able teachers than those who pass the test."

TM 011 199

State Representative Harold Metts, sponsor of Rhode Island Testing Reform legislation.

When state education officials refused to stop requiring the NTE, Rhode Island ACLU Director Steven Brown and Urban League President Joyce Stevos organized a legislative campaign. Their pressure was so effective that both branches of the state legislature overwhelmingly passed the bill and Republican Governor Edward DiPrete signed it over the opposition of state education department leaders.

New R.I. Commission

The Rhode Island Commission will be appointed later this fall. According to the new law, the 11 members must include one representative each from the Rhode Island Urban League, NAACP, and Hispanic Pro-Education Committee as well as three students from racial minority organizations at state colleges. Test specialists from Rhode Island College, the University of Rhode Island and a member of the state Board of Regents for Elementary and Secondary Education complete the commission membership.

New York Reforms

The New York "Item Analysis Study Commission" law signed by Governor Mario Cuomo requires manufacturers of college and university admissions tests to turn over to a 10-member commission correct answer rate data on each question broken down by gender and ethnicity. The Scholastic Aptitude Test (SAT), American Col-

(*continued on page two*)

CONTENTS

Legislative Round Up 2
K-2 Exams Dropped 3
Testing Teachers 4
Test Questions 5
Life Without the SAT 6
Coaching Works 7
Work with FairTest 12
Order Form 15
The FairTest Agenda 16

Legislative Round-Up

Arizona: A new law requires the State Department of Education to publish an annual report on how every school and school district performed on state-mandated multiple-choice tests.

Mississippi: A new law authorizes the State Department of Education 1) to require students to pass a *Functional Literacy Exam* in order to obtain a high school diploma and 2) to administer an unlimited number of additional state-mandated tests in "any grade deemed appropriate." Currently, Mississippi requires basic skills tests in grades 3, 5, 9 and 11. All students in kindergarten, 1st, 4th, and 6th grades take the Stanford Achievement Test.

New York: Two measures were enacted into law. The first establishes a 10-member study commission to examine whether university admission exams are unfairly biased against minorities or women. The second overhauls the state's $40 million "merit" scholarship fund, ending the practice of relying exclusively on ACT or SAT scores to determine awards (*see page 1*).

North Carolina: A new law terminates funding for state-mandated testing of first and second graders (*see page 3*).

North Dakota: A new law authorizes the State Legislative Council to study "the feasibility and desirability" of instituting a program of competency testing of elementary and secondary students.

Rhode Island: A new law creates an 11-member commission to examine whether ETS' National Teachers Exam is unfairly biased against minorities or women (*see page 1*).

Wisconsin: Two bills are pending. The first would extend the Truth-in-Testing concept to all university admissions exams annually administered to more than 500 Wisconsin residents. The second measure would extend the Golden Rule Bias Reduction concept to the Wisconsin Insurance Agent Licensing Exam.

The National Center for Fair & Open Testing (FairTest)
P.O. Box 1272, Harvard Square Station
Cambridge, MA 02238; telephone (617) 864-4810.

Staff: *Jim Gray*, Washington Coordinator; *Noe Medina*, Consultant; *Nancy Murray*, Publications; *Bob Schaeffer*, Communications; *Sarah Stockwell*, Sex Bias Coordinator; *John Weiss*, Executive Director.

We would like to thank the Ford Foundation and the Rockefeller Family Fund for supporting our campaigns to ensure that the standardized tests administered to students and job applicants are fair, open and educationally sound.

States Protect Test-Takers

(from page one)

lege Testing Program Assessment (ACT), Graduate Record Exam (GRE), Law School Admissions Test (LSAT), Medical College Admissions Test (MCAT), and the Graduate Management Admissions Test (GMAT) are covered by the plan. A coalition of civil rights, education, feminist and student groups led by the New York Public Interest Research Group (NYPIRG) campaigned for the new law (see Summer 1987 *FairTest Examiner*).

The change in New York's scholarship program followed release of a joint FairTest/NYPIRG report documenting that more than two-thirds of state scholarships granted solely on the basis of SAT and ACT scores went to males. The State Board of Regents will develop a new award formula which includes high school academic performance.

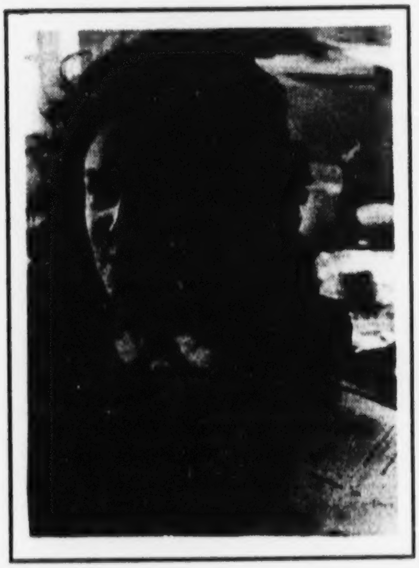
National Implications

The NTE administered in Rhode Island is identical to the teacher tests administered in 10 other states, while the university admission tests being examined in New York State are administered nationwide. Accordingly, the changes recommended by these study commissions will have national implications.

Several other states including Wisconsin and Massachusetts are expected to take up testing reform proposals this fall.

The texts of the new Rhode Island and New York statutes and other key testing reform reports and documents are available in FairTest's recently revised *Standardized Testing Reform Sourcebook, 1987/88*. Copies are $25.00 postpaid. (Use order form on inside back cover.)

North Carolina Legislature Drops Exams for 1st, 2nd Graders



"Legislators are no longer willing to simply accept the promotional words of test publishers"

Page McCullough
Executive Director
of the Atlantic Center
for Research in Education

For the first time since the testing craze swept America's schools a decade ago, a state-mandated testing program has been terminated. Pressured by a coalition of parent, teacher and research groups, the North Carolina legislature has voted not to continue funding CTB/McGraw Hill's *California Achievement Tests* (CAT) for first and second graders.

"This is a tremendous victory, especially for the children," said Page McCullough, Executive Director of the Atlantic Center for Research in Education (ACRE), which has been fighting to reform North Carolina's K-12 testing policy for ten years. "But it is also a victory for everyone who cares about teaching children well."

Legislators deleted the testing program over the strong objections of the State Board of Education, State Department of Public Instruction and former Governor James Hunt. "The questions and concerns raised by legislators during the debate on this issue revealed a level of sophistication about testing that I have never seen before," commented McCullough. "Legislators are no longer willing to simply accept the promotional words of test publishers or the state's testing bureaucrats. Lawmakers have heard too much about the downside from parents and teachers." In addition to ACRE, the victorious coalition included the North Carolina Association for the Education of Young Children, North Carolina for Fair Schools and the North Carolina Association of Educators (NEA affiliate).

"We are overtesting our children in North Carolina. They are very good at filling in answer sheets, but that doesn't teach them to think."

Among the arguments raised by parents and teachers that helped win the debate in the legislature were the following:

- The CAT tests had low *reliability* for young children--students' scores could fluctuate significantly from one test administration to another.
- These tests had low *content validity*--the test did not overlap closely with the state's curriculum requirements.
- Scores from a standardized test provide educators with little useful information because young children develop at very different rates.
- Curriculum was being "test-driven." Teachers were under a great deal of pressure to ensure that their students performed well on these tests. Against their better judgment and the recommended curriculum, teachers were altering what they taught to fit the test.
- Days devoted to testing (and drilling for it) were taking up valuable instructional time.

The state budget now contains a provision banning the purchase of norm-referenced tests in grades one and two and requiring the state to *"provide developmentally appropriate individualized assessment instruments consistent with the Basic Education Plan for first and second graders."* ACRE's McCullough explains: "These are fancy words for individualized non-multiple-choice means of assessing children."

CTB/McGraw Hill reacted to the policy change by offering a 20% discount on the price of the CAT to North Carolina school districts which choose to spend local funds to keep using the tests.

A specially-appointed committee --which will include an ACRE representative--will make sure that the intent of the legislature is carried out. "We are overtesting our children in North Carolina," Karen Garr, President of the North Carolina Association of Educators told the *Raleigh News & Observer*. "They are very good at filling in answer sheets, but that doesn't teach them to think."

For more information about the successful campaign in North Carolina contact: ACRE, 604 W. Chapel Hill Road, Durham, NC 27701, (919) 866-6464.

Testing Teachers...Teacher Testing

St. Louis:

Student Scores No Way To Evaluate Teachers

The St. Louis Board of Education's policy of using student scores on standardized multiple-choice tests to evaluate teachers is unfair because the tests are poorly matched with the district's curriculum, according to an analysis commissioned by the local teachers' union. The study, written by Educational Products Information Exchange of Water Mill, NY, reports that only 9.5% of all content items were common to the test, the district's curriculum guide and assigned textbooks for four grade levels examined. The report has been introduced as evidence in a lawsuit brought by the St. Louis Teachers Union, an affiliate of the American Federation of Teachers, seeking to overturn the controversial evaluation policy.

So far, 22 teachers have been placed on probation largely because their students did not show sufficient gains on the California Achievement Test. While a few other factors were taken into consideration, school officials told *Education Week* (June 3, 1987) that test scores were "a significant, major part of the evaluation." If the current policy remains in place, these teachers are in danger of being dismissed.

> ***"An awful lot of what you teach doesn't get tested, and an awful lot of what's tested doesn't get taught."***
> ***Professor Andrew Porter***
> ***Michigan State University***

Ronald Berk, Professor of Educational Research at Johns Hopkins University, has filed a friend-of-the-court brief in the St. Louis case. He believes that using test scores is not only ineffective in determining teacher accountability, but is also inappropriate and unfair. As he said at a recent AFT conference, "Student achievement tests reflect many variables besides student knowledge or subject comprehension. Socio-economic or emotional factors and intellectual capacity all contribute to how well a student performs. These elements are beyond the parameters of teacher instruction in the classroom."

New Tests?

"An awful lot of what you teach doesn't get tested," Michigan State University Professor Andrew Porter recently told *Education Week*, "and an awful lot of what's tested doesn't get taught." In an effort to sidestep these criticisms, St. Louis School Superintendent Jerome Jones said that changes will soon be made to ensure the tests are better aligned with the curriculum, including the introduction of new tests specifically designed for the St. Louis school system.

Lawsuits Challenge NTE Use in MA, GA

Lawyers representing the Hispanic and Southeast Asian residents of Lowell, Mass. have filed suit in a U.S. District Court charging that a recently adopted desegregation plan fails to address the needs of minority students. The suit charges that the city's requirement that all teachers score above a certain cut-off level on ETS' National Teachers' Exam unfairly discriminates against minority teachers. Roger Rice, an attorney for the plaintiffs, believes that there is a good chance the city will soon drop NTE requirements.

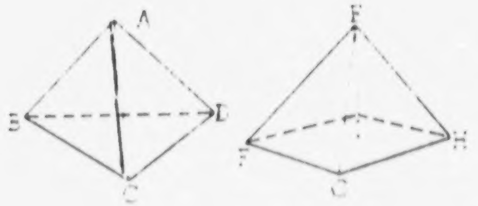
327 experienced teachers in Georgia have lost their jobs and certification because they failed to pass the state's Teacher Competency Test after nine tries. 244 of the teachers are black. The state affiliate of the NEA has filed a suit seeking to prevent Georgia from using the test as a screening device.

17% Fail Tests

Seventeen percent of prospective teachers failed teacher tests in 1986, according to a U.S. Ed. Dept. report, *What's Happening in Teacher Testing?*, released in August 1987. Today, 26 states require prospective teachers to pass a multiple-choice test, and 18 others are planning to do so soon. According to the report, the tests do not "evaluate many important human qualities such as dedication, caring, perseverance, sensitivity and integrity, and they cannot guarantee that an individual who passes will become a good teacher."

Are These Exam Questions Fair, Valid and Educationally Sound?

1)



In pyramids ABCD and EFGHI shown above all faces except base FGHI are equilateral triangles of equal size. If face ABC were placed on face EFG so that the vertices of the triangles coincide, how many exposed faces would the resulting solid have?
a) five b) six c) seven d) eight e) nine

--PSAT

2) **RUNNER:MARATHON**
a) envoy:embassy b) martyr: massacre
c) oarsman:regatta d) referee:tournament
e) horse:stable

--SAT

3) **SOCCER:TEAMS**
a) golf:individuals b) football:spectators
c) badminton:rackets d) coach:players
e) baseball:bases

--SAT

4) **GEARS:MESH**
a) pulleys:lower b) numbers:divide
c) facts:correlate d) files:smooth e) string:entwine

--SAT

5) **Mary had 35 boxes to unpack. She had unpacked 21. What percentage of the unpacking had she done?**
a) 33% b) 40% c) 60% d) 67%

--NJ Proficiency Test for 9th Graders

6) **The city garbage truck has not picked up Esther Maxey's garbage for three weeks. Esther is having trouble keeping the flies and mice away. What should she do?**
a) Take the garbage down the street to an empty lot.
b) Call the hospital to complain about the mice.
c) Call the sanitation department about the problem.
d) Cover the garbage with a sheet.

--ACT's Adult Performance Level Exam

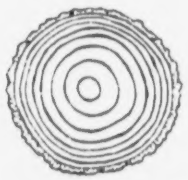
7) **In which continent is Chile?**

--WISC-R IQ Test

8) **What is the capital of Greece?**

--WISC-R IQ Test

9)



Which of the following explains the pattern of rings in the cross section of this tree?
a) A logging operation has hindered its growth.
b) An insect attack destroyed its leaves and reduced its ability to grow.
c) As the tree grew older, it became crowded by nearby trees which had not interfered with its growth earlier.
d) When the tree was young, fire destroyed the leaf covering on the forest floor which had previously held moisture in the soil.

--California Achievement Test

For answers, see page 11.

Rosser to Conduct Sex Bias Research

In the wake of the tremendous public interest in the issue following FairTest's April press conference on sex bias in testing, researcher Phyllis Rosser has received a federal research and development grant from the Women's Educational Equity Program for 1987-88 to continue her investigations.

The project, which will result in a major report and training materials for advocacy work, will focus on three areas:

- assessing the nature and extent of sex bias in the SAT and other exams which have great impact on women's educational opportunities;
- analyzing the impact of girls' lower scores on their receipt of merit scholarships and participation in gifted and talented programs;
- developing policy recommendations and strategies for overcoming sex bias.

Research will include an item analysis of the June, 1986 SAT, administered to students at the Princeton Review, and an analysis of data collected by ETS.

Rosser would like to receive the advice or assistance of anyone in the testing reform network who has studied this issue. Write to:
77 Telegraph Hill, Holmdel, New Jersey 07733.

Sex Bias Report

Sex Bias in College Admissions Tests: Why Women Lose Out, by Phyllis Rosser with the staff of FairTest (April, 1987, 2nd edition) is available for $7.95. Write to FairTest, P.O. Box 1272, Harvard Square Station, Cambridge, MA 02238.

OPINION...OPINION...OPINION

Which Test or Whether to Test

by William R. Mason III

"The real issue is whether any test needs to be required for admission, not which test should be used."

William Mason,
Director of Admissions,
Bowdoin College

Janice Simpson, writing in the April 15, 1987 *Wall Street Journal* suggested that Middlebury College's decision to allow applicants to submit the results from five Achievement tests in place of previously required SATs was "good news for some of next year's applicants." I don't view it as good news at all--for the applicants or the secondary schools that they attend.

The real issue is *whether* any test needs to be required for admission, not *which* test should be used. Having made submission of both SATs *and* Achievements optional back in 1969, Bowdoin College has a wealth of experience in admitting highly-selected classes--3,705 students applied for 385 openings for the class of 1991--without requiring test results. College academic performance of submitters and withholders has been comparable since the policy was begun.

We do not feel that much is gained for the applicant or the admissions staff by shifting the emphasis (over-emphasis?) from SATs to Achievements. Indeed, much may be lost. Such a policy does not lessen the widely held conception of students and parents that the test results dictate the admissions decision. In addition, high school curricula already are shaped to a large degree by the need to excel on standardized tests. The message we should be sending to high schools is that admissions offices at selective colleges are capable of making informed decisions without relying heavily or at all on the Educational Testing Service, not that we want them to design their courses to what can be tested by multiple-choice exams. By and large, that isn't the way we evaluate our students. Why should we require that high school students be evaluated in that fashion?

Our position has been consistent from day one and it has not been anti-test. We find that test scores can be helpful, and we even use them for counseling purposes once the class is matriculated. But do we need them? The answer is a resounding "no." In a climate where parents, guidance counselors, school boards, and headline writers all overemphasize the importance of test scores, that is the good news.

Studies Prove Coaching Works

Can short-term preparation boost students' scores on the SAT and other standardized tests? For years ETS and the College Board have told students that "the only 'preparation' for the SAT that might have an effect on the score is, in a sense, a lifetime of intellectual challenge." If, on the other hand, coaching works, then the tests are measuring students' test-taking ability, and not their academic potential. In addition, low income students who cannot afford high-priced coaching courses will be further penalized by the exam. As College Board consultant Dean Whitla of Harvard has written, "If the Board's tests can regularly be beaten through coaching, then the Board itself is discredited."

Coaching Studies

Three recent studies of the effectiveness of coaching for the SAT report large gains for students. In 1984, Howard University Professor Sylvia Johnson examined the impact of a 30-hour NAACP sponsored SAT preparation course specifically designed for minority youths. She found that the average SAT scores of students increased from 36 to 87 points, with average verbal increases ranging from 11 to 49 points.

"Excessive Coaching"

The principal, another administrator and three teachers at a Fairfax County, Virginia elementary school have been disciplined because of "excessive coaching" for the SRA test given to all 4th, 6th, 8th and 11th graders in the county. Officials claim that coaching at Bailey's Elementary School went beyond permitted lessons in test-taking skills and may have included the use of last year's test questions for practice.

This year, Gregory Coffin, Director of the Urban Schools Collaborative at Northeastern University, reported on the results of a study using computers and instructors to prepare disadvantaged students to take the SAT. His 25-hour course resulted in 91-point gains for students (51 points verbal, 40 points math). Coffin told the May 20, 1987 *Education Computer News* : "By teaching test-wiseness to those who score low on the socio-economic ladder, we attack the equity issue."

"If the Board's tests can regularly be beaten through coaching, then the Board itself is discredited."

Zuman Study: 110 Point Gains

Harvard University Doctoral Candidate John Zuman recently published the results of his controlled study of 88 students taking the Princeton Review's SAT prep course, *The Effectiveness of Special Preparation for the SAT: An Evaluation of a Commercial Coaching School.* Two groups of New York City high school students participated: 55 regular Princeton Review clients who paid $595 to enroll in the course and 33 low-income minority students attending on scholarships.

Students were tested before and after they took the course. Zuman reports that regular students averaged a 110-point increase (+58 math points, +52 verbal points). On average, scholarship students gained between 57 to 74 points on the math SAT, but showed no improvement on the verbal part of the exam.

Zuman says that many scholarship students came from families where English was their second language and that they did not therefore have the necessary vocabulary to deal with SAT-type verbal questions. He also thinks that some of the scholarship students were confused by the Princeton Review's verbal material, and thus experienced a decrease in their verbal scores.

The results of the Johnson, Coffin and Zuman studies support the conclusion that intensive, well-planned courses focused precisely on SAT-type questions *can* have a significant effect on increasing students' SAT scores. As Zuman observes: "The question is no longer whether coaching can be effective, but what makes coaching effective."

Zuman also recognizes the implications of his research: "If coached students have inflated SAT scores, then colleges may well be turning away some more talented students only because those students did not attend an effective SAT preparation course, and may instead be admitting students from wealthier homes."

Research to Examine Computerized Coaching

A doctoral student at Boston University will be evaluating two commercial software packages--one produced by Harcourt Brace Jovanovich, Inc. and the other by Hayden (Spinnaker) Software. The dissertation will examine whether use of these products reduces anxiety and increases scores. A pilot study was conducted with 40 students last May. For more information, contact Fr. Richard Curran, 637 Cambridge St., Brighton, MA 02135, (617) 783-9090.

Bennett Pushes NAEP Expansion

Secretary of Education William Bennett and Senator Edward Kennedy have introduced separate proposals to expand the federal government's National Assessment of Educational Progress (NAEP) testing program from under $4 million to more than $26 million a year. Under both Bennett's and Kennedy's proposals, states will be compared by their students' scores on NAEP tests. The Senator's proposal would make participation in the program optional, while Sec. Bennett would require states to administer the NAEP exams. Currently NAEP is run by the Educational Testing Service, which is lobbying hard for an expansion of the program.

For further reading on NAEP, see "Our NAEP scores are better than your NAEP scores!" in *R&D Preview*, July 1987, obtainable from the Council of Educational Development & Research, Suite 304, 1201 16th Street, NW, Washington, DC 20036.

Also see *The Nation's Report Card: Improving the Assessment of Student Achievement*, available from the Harvard Graduate School of Education, 108 Longfellow Hall, Cambridge, MA 02138 for $9 a copy.

UCLA Researchers Find Tests May Be Unfair

The Center for the Study of Evaluation at the University of California at Los Angeles has conducted a national survey on testing programs in public schools.

Their 190-page report concludes that standardized competency tests "may be unfair for many students. They may be reducing the frequency of routine classroom testing and (thus) the quality of instruction. They may be narrowing the curriculum and, with it, the range of opportunities open to many students. These possibilities deserve the attention and investigation of all those who shape educational policy at the local, state, and national levels."

Copies of *Assessing Student Achievement: A Profile of Classroom Practices* by Donald W. Dorr-Bremme and Joan L. Herman are available from the Center for the Study of Evaluation, Graduate School of Education, University of California, Los Angeles, CA 90024 for $6 + $1.50 handling and postage.

Congressional Report Says Tests Poor Measure of Reform

An August, 1987 report, prepared for the Senate Education, Arts and Humanities Subcommittee by the Congressional Budget Office, argues that the recent rise in the scores of students on standardized tests is not attributable to the decade's school reforms. Contending that testing trends are an inadequate yardstick for measuring reform--which hinges on a far more complex range of factors than tests can measure--the report also claims that the growing tendency by teachers to "teach to the test" could skew the significance of test scores.

Copies of the 102-page report, *Educational Achievement: Explanations and Implications of Recent Trends*, may be obtained at no cost by writing to the Publications Office, Congressional Budget Office, U.S. Congress, Washington, DC 20515.

N.Y. School Shut for Low Test Scores

New York City Schools Chancellor Nathan Quinones has ordered the closing of Bronx Junior High School #123 because reading test scores failed to rise. According to the *New York Times* of August 10, 1987, Quinones said he would place "considerably more than five schools" on notice for the 1987-88 academic year. Unless these schools boast their average test scores, Quinones warns, they too will be closed.

More Drop Outs?

Starting next year, students in grades 1, 5, 8 and 12 in the Boston public schools must pass a standardized multiple-choice test in order to be promoted.

If the testing plan had been in place last year, approximately 40% of high school students, 32% of the middle grades students, and 30% of the elementary students would not have been promoted, according to the Massachusetts Advocacy Center. The Center cited research showing that when a student repeats one grade, dropout chances increase by more than 40%.

For more information, contact Anne Wheelock, Mass. Advocacy Center, 76 Summer Street, Boston, MA 02210, (617) 357-8431.

> "Teachers will teach whatever is required to enable their students to do well on assessment tests, and textbooks will be written to match the test objectives. Too often today's assessments dictate a curriculum filled with rote calculation and mimicry mathematics."
>
> --Prof. Lynn Arthur Steen
> *Science Magazine*,
> July 7, 1987

N.Y. Legislature Votes to Require Testers to Divulge Details of Students' Responses

Critics of standardized tests hail measure, confident they will find evidence to back their bias claims

ALBANY, N.Y.

In a major victory for critics of standardized tests, the New York Legislature last week voted to require testing companies to give the state detailed information about the performance of various types of students on their examinations. Companies would be required to provide a breakdown of the rates at which correct answers were provided on each question by men and women, and by students of different races and ethnic and linguistic backgrounds.

Gov. Mario M. Cuomo, a Democrat, is expected to sign the legislation, which testing critics have said would provide them with their first-ever opportunity to gather the kind of evidence needed to prove that certain questions used on standardized examinations, such as the Scholastic Aptitude Test, are biased against members of certain groups.

That evidence, they said, would allow them to push for further legislation in New York and elsewhere to establish procedures to force testing companies to discard such questions.

'We're Going to Get the Proof'

"For almost 20 years, people have been saying the tests are biased, but we've had only total scores to go on," said John G. Weiss, executive director of the National Center for Fair and Open Testing. "Now we're going to get the proof we need."

The College Board, which sponsors the S.A.T., and the Educational Testing Service, which administers it, have both denied that their standardized tests are biased against women and minorities.

In addition to requiring that testing companies file reports on the correct-answer rates of the various groups during a one-year period, the New York Legislature also voted to create a special committee to study, based on the reports of the testing companies, whether the tests are biased and should be changed.

Although the study itself would be limited to this state, actions against the testing companies here have previously led to national changes in the way the standardized tests are conducted.

In 1979, New York adopted the first "truth-in-testing" legislation, requiring the testing companies to provide students with their corrected answer sheets and a copy of the examination. E.T.S. extended those rights to the rest of the country.

During the past two years, testing critics have renewed their efforts here and elsewhere to broaden "truth-in-testing" laws.

Mr. Weiss said the New York legislation would help testing critics push for adoption

Continued on Page 22, Column 2

N.Y. Votes to Require Testers to Divulge Information

Continued from Page 17

by the testing companies of the "Golden Rule" procedure, which settled a lawsuit between the Golden Rule Insurance Company and the Educational Testing Service. That procedure calls on the companies to separate all test questions into two categories: those on which the correct-answer rates of blacks and whites (or women and men) differ by less than 15 percentage points and those that differ by more than 15 percentage points. Questions in the first category would then have preference for use on future exams.

Testing companies have opposed further use of the Golden Rule procedure, instead proposing the use of methods that would compare the correct-answer rates on individual questions of black and white students who score at the same overall levels.

The approach favored by the testing companies is called the Mantel-Haenszel method, after the two researchers who formulated it.

After the legislation to create the study committee was amended to require the panel to study both methods for removing bias, the major testing companies withdrew their opposition to the bill.

Statement by College Board

The College Board said in a statement that it had never taken a position on the bill and urged the study committee to "separate objective evidence from subjective inference."

Last year Frederick H. Dietrich, vice-president for programs of the College Board, testified against a version of the bill at a legislative hearing here. He said he opposed the use of the correct-answer rates of different ethnic groups on questions to study alleged bias on the tests.

"Since performance among racial/ethnic groups may reflect educational disadvantages, any attempt to 'remedy' those disadvantages by manipulating the tests, rather than by addressing the substantive issues of finding ways to enhance educational preparation of all students, would be both misleading and harmful to the very students we believe the sponsors of this bill have at heart," Mr. Dietrich said.

Asked about Mr. Dietrich's statement, a spokesman for the College Board called it "informational testimony," and insisted that the board had never opposed the legislation.

Alice J. Irby, vice-president for field services of the Educational Testing Service, said, "E.T.S. has never opposed legislatures studying things," but acknowledged that the service had opposed the bill previously "because of concerns about fairness."

In the bill's present form, with the specific requirement that the committee study the Mantel-Haenszel method, Ms. Irby said, "there is a reasonable prospect that a fair study can be conducted."

An aide to Kenneth P. LaValle, chairman of the New York Senate Higher Education Committee and sponsor of the legislation, said the charge to study the Mantel-Haenszel method would probably bolster the position of the testing critics. "Now no one will be able to say, 'You didn't study the thing that would disprove your theory,' " she said.

—SCOTT JASCHIK

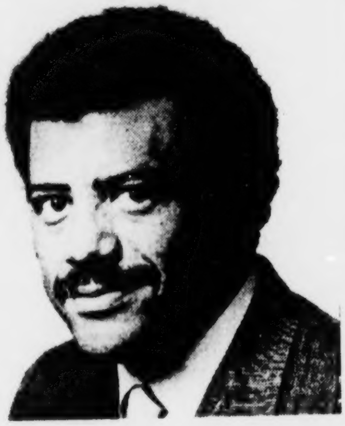
June 24, 1987 • THE CHRONICLE OF HIGHER EDUCATION •

In the News

John J. Fremer has left his position as Vice President of the Psychological Corporation (a subsidiary of Harcourt Brace) to become Senior Development Leader at the Educational Testing Service. His assignment is to develop a new generation of tests and other assessment devices for ETS. Fremer is co-chair of the American Psychological Association's committee that is developing a *Code of Fair Testing Practices in Education.*

Diana Pullin has left Michigan State University to become the new Dean of Boston College's School of Education.

E.D. Hirsch, Jr., author of the bestselling book *Cultural Literacy: What Every American Needs to Know*, has established a foundation to develop and market a standardized test of students' general knowledge.



Bernard Gifford, Head of Commission

Dean of the School of Education at the University of California at Berkeley *Bernard Gifford* has been awarded an $800,000 grant from the Ford Foundation to establish a *Commission on Testing & Public Policy*. The three-year investigation will examine the "role of standardized paper and pencil testing in the allocation of opportunities in the U.S. today." The 17-member Commission will be announced in November, 1987.

Worth Reading

- "Questions to Ask When Evaluating a High Stakes Testing Program," by Diana Pullin and George Madaus, *NCAS Backgrounder*, June 1986. From NCAS, 100 Boylston Street, Suite 737, Boston, MA 02115.
- The State Higher Education Executive Officers Association recently recommended that states should publicize the performance of students at individual colleges on licensure and certification examinations. Copies of the SHEEO report are available from SHEEO, 1860 Lincoln Street, Denver, CO 80295.
- *Golden Rule or Ruse* is the theme of a special Summer 1987 issue of *Educational Measurement*. Articles by Golden Rule President Patrick Rooney, ETS Researcher Jane Faggen, Pittsburgh Professor Lloyd Bond and FairTest Executive Director John Weiss. For $5.00 from NCME, 1230 17th Street, NW, Washington, DC 20036.

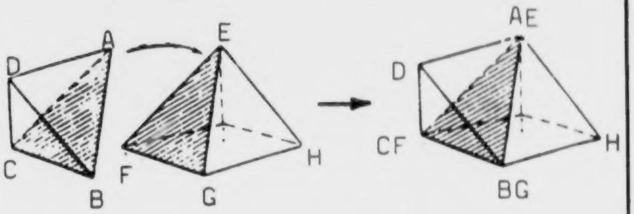
Upcoming

- October 14-17: American Evaluation Association annual conference in Boston. For more information, call Sandy Sherman (804) 924-0511.
- October 28-30: College Board Annual Convention, Chicago Marriott, Chicago, IL (312) 713-8000.
- October 31: ETS' 1987 Invitational Conference at New York City's Plaza Hotel. For more information, contact M. Lamb, ETS, Princeton, NJ 08541.
- November 12: *Testing Bias: Reverse Discrimination & Affirmative Action* will be the theme of Fair-Test Executive Director John Weiss' presentation at the Annual Meeting of the Rhode Island Civil Liberties Union. For more information, call (401) 831-7171.
- November 18-20: The Third Biennial Conference on Minority Assessment will meet concurrently with the 18th Annual Conference of the Rocky Mountain Educational Research Association in Tucson, Arizona. The sponsors are currently inviting proposals for papers. For further information, contact Karen Gendron, Conference Coordinator, Education Building, Box 604, University of Arizona, Tucson, Arizona 85721.
- April 5-9, 1988: The American Educational Research Association and National Council for Measurement in Education will hold their 1988 annual meeting in New Orleans, LA. Address inquiries and send proposals for papers to the AERA Meeting Coordinator, 1230 17th St., NW, Washington, DC 20036.

Answers to Test Questions

(see page 5)

1. **(A) five.** ETS and their panel of 16 experts selected *c) seven* as the wanted answer. High school student Daniel Lowen used New York's Truth-in-Testing law to challenge their answer successfully. He showed that when the pyramids were joined, four of the original triangular faces were merged into two solid planes (*see below*). As a result of his challenge, 200 additional students became National Merit Scholarship Semi-Finalists.



2. **(C) oarsman:regatta.** This question is typical of many SAT items that require students to be familiar with the activities and vocabulary of America's upper middle class. A partial analysis of this question made public by ETS indicates it was of average difficulty for whites (53% answered correctly), but very difficult for blacks (only 22% gave the "wanted" answer). If the Golden Rule procedure (*see page 12*) were applied to the SAT, this item would be replaced by another analogy item with approximately the same correct answer rate but which did not require knowledge of such culturally-specific terms as 'regatta'.

3. **(A) golf:individuals.** ETS researcher Alicia Schmitt found that Hispanic test takers scored significantly lower than expected on this item. Might it be because fewer Hispanic high school students were knowledgeable about golf? Dr. Schmitt notes: "Items with content reference of special interest for Hispanics seem to help performance." For example, Mexican-American students scored signficantly higher than expected on a reading comprehension passage concerned with lifestyle changes in Mexican-American families. See Alicia Schmitt, *Unexpected Differential Item Performance of Hispanic Examinees on the SAT-Verbal*, Unpublished ETS Statistical Report, 1986.

4. **(C) facts:correlate.** 46% of the males, but just 27% of the females marked the wanted answer . Such a wide difference in scores is almost certainly attributable to the male-oriented subject matter of the question.

5. **(C) 60%.** CTB/McGraw Hill said this was the "wanted" response: 21 divided by 35. New Jersey school teacher Jeff May challenged their answer. He argued that the word "had" in the question could mean that Mary had already unpacked 21 boxes but still had 35 left. In that case the correct answer would be 37.5%. Students who calculated the problem this way would then mark *(B) 40%*, since they are often asked to estimate answers, and pick the closest one.

6. We suppose the correct answer is **(C) call the sanitation department about the problem**. But given the inefficiency of many public services, a plausible argument could be made for any of the answers. Walt Haney and Lloyd David write that "if this item measures anything at all, surely it is the test-taker's ability to ferret out the item writer's sense of social propriety. But the notion that this sort of test item measures 'adult competence' or even competence in figuring out what to do with one's garbage is altogether preposterous." See Haney and David in *APL Revisited: Its Uses and Adaption in States*, National Institute of Education, 1980.

7. **South America**
and
8. **Athens**. These two questions are just a few of the WISC-R I.Q. items that require children to be familiar with distant places. Thus, children who travel frequently or whose families are more attuned to international affairs have a greater tendency to answer such items correctly. Due to a lawsuit initiated by the Association of Black Psychologists, the WISC-R test may not be administered to black children in California where it had been used to track minorities into "dead end" Educably Mentally Retarded classes (see Spring 1987 *FairTest Examiner*).

9. Jeff May also challenged this item arguing that there are two possible right answers: (B) or (C). Maine Audubon Society researcher Melissa Lee agrees: "Either an insect attack or overcrowding from nearby trees could stunt a tree's growth."

FairTest Agenda

(continued from back page)

groups. Independent researchers can analyze this information to determine if test items are measuring extraneous or irrelevant information. Legislators in New York and Rhode Island have established item analysis study commissions to determine whether tests administered in their states are fair (*see page 1*).

GOLDEN-RULE BIAS REDUCTION PRINCIPLE: This reform is modeled after the landmark 1984 out-of-court agreement ending a lawsuit by the Golden Rule Insurance Company against the Educational Testing Service. The insurance company had charged that ETS' Illinois Insurance Agent Licensing Test was racially biased and not job related.

The test construction techniques agreed to in the settlement make *exams fairer, but not easier*. Using the Golden Rule Principle, test-makers cover the same content areas as previous tests, and the test remains at the same level of difficulty.

The only difference is that within groups of items measuring the same content area, test-makers must select those items on which minority and majority group candidates differ least. This ensures that questions measure genuine knowledge differences and not irrelevant, often culturally specific, information.

As a result of the 1984 settlement in Illinois, the gap between black and white test-takers on the Illinois Insurance Test has been reduced by more than a third. The Golden Rule Principle has been incorporated in the settlement of a teacher testing lawsuit in Alabama and in legislation currently pending in several states.

Work with FairTest

MANAGING DIRECTOR

Duties: Responsible for overseeing day-to-day implementation of all FairTest projects including public education, advocacy, publications and fundraising. Work with Executive Director to design strategy and allocate resources.

Requirements: Strong staff supervision, planning and fundraising background; experience working on civil rights and/or education issues a plus.

Salary: Up to low 30s, health care and other benefits.

PUBLICATIONS/ FUNDRAISING COORDINATOR

Duties: Publish quarterly newsletter, topical updates and special reports. Assist with fundraising campaigns including proposal production, list acquisition and donor identification.

Requirements: Strong editing, writing and research skills; interest in education and civil rights issues.

Salary: Competitive, with benefits.

To apply, send us a cover letter and resume as soon as possible.

Information Sought on Test Scores and Scholarships

FairTest is hoping that readers can assist in our on-going research projects.

- Do you know of scholarships that are dependent on test scores?
- Do you have data on the gender/race breakdowns of scholarship recipients?
- Can you identify students who have gained admission to a college or graduate school but have been denied financial aid exclusively because their test scores fell below a certain level?

We are also interested in race/gender information for all college and graduate school admissions and professional licensing exams. If you can help us, please write to FairTest, P.O. Box 1272, Harvard Square Station, Cambridge, MA 02238.

Standardized Testing Reform Sourcebook

(1987-1988)

Completely revised edition of FairTest's popular looseleaf notebook resource for advocates, lawyers and educators. Includes key legislation, court decisions, reports and academic articles covering the Golden Rule Principle, Truth-in-Testing and Item Analysis. An essential reference for anyone working on testing reform.

$25.00 Postpaid
(approx. 200 pages)

use order form on page 15

As reliance on testing increases, so do charges of test bias

Along with education reforms, school districts nationwide have seen an increased emphasis on standardized tests to measure performance and increase accountability.

With more testing, concerns have grown about tests' relevance and accuracy, leading in some cases to court challenges and proposals for new legislation.

The New York state legislature, for example, recently approved a bill to initiate a study of standardized tests required for college admission to determine if the tests are biased against women and minorities. Gov. Mario Cuomo is expected to sign the bill.

A related bill would change the policy of awarding New York state scholarships solely on the basis of Scholastic Achievement Test (SAT) scores.

Some legislators have suggested broadening the criteria to include grades, teacher recommendations and class rank, but no consensus has been reached and the legislative clock may run out.

The second bill was introduced following a study released in March by the New York Public Interest Research Group (NYPIRG) reporting that boys received more than 70 percent of the 1,000 Empire State scholarships awarded in 1986-87.

According to Blair Horner of NYPIRG, the SAT unfairly favors white males, producing a 61-point gap between male and female scores. Yet, girls typically have higher grade point averages in both high school and college.

Jeffrey Bowen of the New York State School Boards Association said figures released by the state Education Department show a similar disparity for the state Regents College scholarships awarded this year. Only for state nursing education scholarships do female recipients far outnumber males. Both of these scholarships also are awarded according to SAT scores.

Bowen said that where local school districts are concerned, "the potential major impact I would see at this point, if in fact the criteria was broadened, would be additional paperwork."

Disparities in testing experiences between whites and minorities also have garnered attention in some states, including Texas.

"The focus on testing in Texas since the 1984 reform movement has been greater than ever," according to Doug Rogers of the Texas Association of School Boards, and "some problems have been encountered."

Two bills concerned with fair testing were introduced during the past session of the Texas legislature, which ended June 1.

One bill would have increased reporting requirments for test manufacturers. The other would have addressed possible racial and ethnic bias in the pre-professional study test (PPST) given to students applying to teacher education programs. It would have required testmakers to throw out any questions that showed more than a 15 percent difference in the rate of correct answers between racial groups.

The bill also would have prohibited educational institutions from denying applicants admission solely on the basis of test scores.

Although both bills died in committee, the Texas Education Agency plans to have a new test developed and currently is weighing several proposals, said Nolan Wood, director of the agency's division of teacher assessment. Unlike the norm-referenced PPST, developed by the Educational Testing Service, the new test will be criterion-referenced.

Wood said the state legislature voted in 1981 to have a test developed because in Texas, "we were really concerned that students didn't have the basic skills" they needed for teacher education.

On the other hand, Texas "has a large minority population—we need Hispanic and black teachers," Wood said, and the PPST "is just running them off."

Since the test was implemented in 1984, the percentage of whites passing has hovered consistently in the upper 70s. The percentage of Hispanics who pass the PPST has increased from 28 percent to 48 percent. For black test-takers the pass rate has increased from 12 percent to 28 percent, but, Wood noted, fewer and fewer black students are even taking the test.

Because of these disparities, civil rights groups representing blacks and Mexican-Americans have been pursuing a lawsuit against the agency. The case is on hold since an appeals court found in favor of the plaintiffs and remanded the case to the district court. The plaintiffs have taken no further action because the agency decided to rewrite the test, Wood said.

The new test, to be introduced in either fall 1988 or spring 1989, will not "water down" the qualifications the state seeks in its future teachers, Wood said. Rather, the state believes a criterion-referenced test will provide a more accurate measure of applicants' abilities and potential.

A connection between teaching and testing also has led to a lawsuit in St. Louis, Mo. The local teachers union filed a lawsuit disputing the district's decision two years ago to link teacher evaluations to students' performance on the California Achievement Test.

The teachers group contends that the policy is unfair because the test is not consistent with the district's curriculum. The union recently commissioned a study by an independent research group whose findings corroborated this argument.

"We think we presented the court with convincing evidence that there is a very good fit between the test and the curriculum," said Norman Walsh, associate superintendent. The school district has administered the test for about seven years, he said.

The district's decision to link teacher performance to student test scores "is just a reflection of the fact that we feel the school system has to be accountable for student learning," Walsh explained. He also noted that test scores are not the only factor used to evaluate teachers.

The St. Louis school district is developing its own standardized tests as part of its comprehensive assessment plan, Walsh said.

The school district also will use the locally developed tests to meet new state test requirements. Starting next year, Missouri will require districts to test third, sixth, eighth and 10th graders in reading, science, social studies and math.

Districts do not have to use state-developed tests if they present evidence that their tests measure the same skills.

"We started our test development about three years ago, so we intend to use our tests," Walsh said.

"I really think that the technology of testing is very well established," he said. "I see no reason why local school districts, if they employ talented people, shouldn't be as capable, if not more capable, of developing reliable tests to evaluate student progress and district objectives."

In many states, how and which tests are administered "is an incredible patchwork quilt," said Bob Schaeffer of FairTest, a Boston-based test watchdog group. States typically require districts to test students at several benchmark points, but testing programs "in many cases are put together at the district level," he said.

More testing means more tests, and school officials should examine new products carefully, Schaeffer said. FairTest suggests three basic criteria: tests should be fair and unbiased, open to investigation and educationally sound.

FairTest publishes materials on standardized testing. For information, write FairTest, P.O. Box 1272, Cambridge, Mass. 02238; (617) 864-4810.

SCHOOL BOARD NEWS
For Local Government of Public Education
July 8, 1987
Published by the National School Boards Association
Vol. 7, No. 12

Some Colleges Rethinking Value of SATs

Administrators Say Test is Taking on Undue Significance With High School Students

Reprinted from The Washington Post, *June 24, 1987. By Staff Writer Barbara Vobejda.*

The Scholastic Aptitude Test (SAT)--a cause of perennial teen-age anxiety and a staple in the college admissions routine for five decades--is coming under new scrutiny in higher-education circles at the same time high schools and their students are assigning greater significance to this academic rite of passage.

Vermont's elite Middlebury College announced in April that it would no longer require applicants to take the SATs. Union College, a New York liberal arts institution that had required taking the test for more than 50 years, made a similar announcement a few weeks earlier. Bates College in Maine declared the SAT optional three years ago. And Bowdoin College hasn't required scores since 1969.

College officials have raised several concerns about the test: It does not predict college performance as well as high school grades or subject-area achievement tests; low scores can unnecessarily discourage students from applying to certain schools; and minorities, women and students from low-income families consistently score below white males.

But many administrators said they were most concerned that the test has taken on undue significance among high school students and their parents. The result is high stress and increasingly popular, high-priced test-preparation courses that can give unfair advantage to students from middle- and upper-income families.

"We just got very dismayed at the amount of emotional energy being devoted to SATs by students and their parents," said Peter Blankman, spokesman for Union College. "It's almost become the tail that wags the educational dog."

The actions of a few schools can hardly be classified as a stampede--1,600 institutions still use the SAT, and in the last decade hundreds of institutions have begun to require the test. But decisions by these selective, Eastern colleges carry weight around the country. The move away from SATs also indicates that the long-standing internal deliberations among administrators about the pros and cons of standardized testing have intensified.

Several other voices are contributing to the debate. Students at Brown University organized and came close to passing a referendum in April calling for administrators to declare the test optional. The students, who vowed to organize similar protests on other campuses, claimed that the test is culturally biased and flawed as a predictor of scholastic success.

Also this spring, the National Center for Fair and Open Testing, a Cambridge-based public-interest group, released a study charging that women were losing millions of dollars in scholarships because they score an average of 61 points below men on the SAT.

Finally, in a report issued this year, the Carnegie Foundation for the Advancement of Teaching urged institutions to drop standardized test scores as admissions requirements unless they are used as a yardstick in the selection process. Ernest L. Boyer, who heads the foundation, said two-thirds of the 30 institutions he studied rely only marginally on the scores, but continue to require them to maintain "an aura of selectivity."...

The College Board, which administers the test through the Educational Testing Service, agrees that high school grades are a better predictor of academic success, but argues that the prediction is even more accurate when SAT scores are combined with grades. The organization also offers 14 achievement tests designed to measure a student's knowledge of a particular subject such as mathematics, history or a language....

The SAT, which was taken by nearly 1.7 million students last year...was created in 1926, when the notion of measuring basic intelligence, rather than achievement, was in vogue. SATs grew to prominence in the postwar years as higher education boomed and college administrators sought ways to assess the wave of prospective students.

The standardized nature of the test was an attempt to compensate for great variations in the quality of high schools and teachers, a point that takes on some irony in the current era when high schools, districts and states are intent on comparing themselves based on their students' SAT scores.

"Now, 60 years later, we're less confident we can measure innate ability so we've started to use the score to measure school performance, absolutely turning the intention of that examination on end," Boyer said.

The SAT is not the only standardized test whose status is changing. Last year, Johns Hopkins University dropped the requirement that applicants to its medical school submit scores for the widely used Medical College Admissions Test. The decision, said spokeswoman Ghita Levine, was made so that undergraduates could spend less time preparing for the test and "to encourage a broader type of student to consider medicine as a career."

In the following year, the applicant pool increased by 50 percent, she said.

Harvard's business school announced two years ago that it would no longer require students to take the Graduate Management Admissions Test, pointing again to an overemphasis on scores and the failure of the test to predict which students would make good general managers.

But the SAT has received the greatest scrutiny because "it is such a symbol of selective admissions," said James Crouse, a professor of education and sociology at the University of Delaware and coauthor of a book on the test.

"There's a growing group of people in positions to influence policy-makers who are beginning to question the test, largely on the basis of academic and scholarly studies," he said.

But others, even those at colleges where the SAT is optional, said their concern is motivated less by problems with the SAT than by the over-reaction it has prompted in high school circles.

"We're not doing it to hammer the SAT," said Rick Dalton, director of enrollment planning at Middlebury. "We're doing it to make a statement to kids. They'll get in because of what they do well, not because of what they don't do well."

At Middlebury, students can still submit SAT scores along with scores from three achievement tests. They also have the option of submitting results from the American College Testing Program, known as the ACT, or five achievement tests.

William Hiss, dean of admissions and financial aid at Bates, said the college's 3-year-old policy making SATs optional has been "wonderful...partly for the effect on our applicant pool. We really have attracted applications from rural Mainers and minority kids."

Roughly one-quarter of the college's freshman class of 400 now enrolls without submitting SATs, he said, relying instead on achievement tests. He said studies conducted over the past eight years have convinced administrators that they have lost virtually no ability to predict a student's success at the school.

More important, he said, is that the college wanted to encourage high school students to refocus their energy.

"We wanted youngsters, instead of spending 400 hours studying word analogies...to take advanced placement calculus, to read every thing Thomas Hardy wrote, to captain the soccer team...to work in a soup kitchen," he said, "to do something that seemed more valuable than simply studying word analogies."

Subscribe to the *FairTest Examiner*

The *FairTest Examiner* will keep you up-to-date on the growing movement for fair & open testing. You will be able to do more than just read about FairTest campaigns--you can help make them successful!

Subscribe today: only $15

Enclosed is my check to FairTest for $_____ to cover the following publications shipped to me postpaid.

____ ***FairTest Examiner***, published quarterly...$15/year
____ ***Standardized Testing Reform Sourcebook, 1987-88*** (200 pages)...$25
____ ***Sex Bias in College Admissions Tests: Why Women Lose Out***, by Phyllis Rosser and FairTest staff, 2nd edition (40 pages)...$7.95
____ ***The Reign of ETS***, by Allan Nairn and Ralph Nader (550 pages)...$30
FairTest Updates issued bi-annually: Spring & Fall. $7/year per topic:
  ____ *Sex Bias in Testing*
  ____ *Golden Rule Reform*
  ____ *Test Coaching*
  ____ *Truth-in-Testing Legislation*
  ____ *Teacher Testing*
  ____ *University Admissions Tests*
  ____ *K-12th Grade Testing*
____ *Spring and Summer* **FairTest Examiner** *$4.00 each (circle issue).*
____ **TOTAL PAID**

Name ______________________ Title ______________________

Organization ______________________

Address ______________________

______________________ Zip ______________________

Please make checks payable to FairTest and mail to:
FairTest, P.O. Box 1272, Harvard Square Station, Cambridge, MA 02238.

THE FAIRTEST AGENDA

Each month hundreds of people subscribe to *The Examiner* or ask to receive FairTest literature. Most support testing reform, but many are unfamiliar with FairTest's background and the terminology of current testing reform proposals. Here are some answers to the questions we most frequently receive.

Who We Are

FairTest is a national, non-profit research and advocacy organization working to eliminate the abuses and misuses of the millions of standardized multiple-choice tests annually administered to America's students and job applicants. FairTest believes that standardized tests often serve as unfair racial, class and cultural barriers to full access to our nation's institutions and that America's overreliance on standardized multiple-choice tests is forcing educators to teach to the limited content of what such tests can measure.

What We Support

FairTest works to ensure that Americans are evaluated in a manner which is fair, valid, open and educationally sound. Tests should do what their promoters claim they will do: they should accurately measure knowledge, skills or experience and predict future performance with a high degree of confidence. Test results must not be influenced by extraneous racial, class or culturally-specific characteristics.

Test-takers and independent researchers need access to old test forms and statistical data on group performance to verify test publishers' claims concerning the fairness and validity of their products. Legislators, educators, parents and students should place scores in their proper perspective, ending overreliance and misinterpretation. Standardized exams should never be used as the single measure of individual or group ability or be allowed to serve as the sole criterion for admission, licensing or promotion.

Proposals for Reform

TRUTH-IN-TESTING: Test publishers should make public old forms of their exams and allow test-takers to obtain copies of their answer sheets and the "correct" answer key. Such provisions, known as Truth-in-Testing, enable test-takers to learn from their mistakes, give future test-takers the ability to better prepare for the exam and give independent researchers the opportunity to examine test questions for fairness and accuracy.

Truth-in-Testing procedures now cover most college and university admissions exams as a result of legislative action in the early 1980s. More than a half-million student scores have been raised due to errors identified under these provisions. Unfortunately, today neither state-mandated tests used in grades K-12 nor certification and licensing exams required to gain employment are covered by Truth-in-Testing safeguards.

ITEM-ANALYSIS: Item-analysis requires test manufacturers to make public correct answer rates on each test item for ethnic and gender

(continued on page 12)

FairTest
National Center for Fair & Open Testing
P.O. Box 1272
Harvard Square Station
Cambridge, MA 02238
(617) 864-4810

Bulk Rate
U.S. Postage
PAID
Cambridge, MA
Permit No. 53278

Complimentary Copy.
Please subscribe today.
See page 15.

Address Correction Requested

FairTest

EXAMINER

The National Center for Fair & Open Testing Volume 2, Number 4 Winter 1988

Beyond Standardized Tests:

New Report Endorses Admissions Alternatives

Students, faculty and alumni at seven highly-competitive colleges and graduate schools which recently dropped or de-emphasized standardized admissions tests strongly praise their institutions' reforms, according to a new FairTest report.

Beyond Standardized Tests: Admissions Alternatives That Work, written by FairTest researcher Amy Allina, documents the experiences of Bates, Bowdoin, Union and Middlebury Colleges, the Massachusetts Institute of Technology, Harvard Business School and Johns Hopkins Medical School. The report also lists three dozen other colleges, including such selective schools as Antioch, Bard and Hampshire, which do not require applicants to take the Scholastic Aptitude Test (SAT) or the American College Test Program Assessment (ACT).

Amy Allina

Reduce Biased Impact

Unfair biases against women and minorities, distortions due to coaching and the impact of tests on curriculum and self-esteem were among the concerns cited by schools which de-emphasized test scores. Typical of reactions reported by FairTest was MIT Admissions Director Michael Behnke's recent testimony before a

(continued on page two)

Teacher Test Flunks New York Study: Movement to Overhaul Exams Grows

The NTE (formerly known as the National Teachers Examination), the country's most popular test for licensing classroom educators, is tainted by gender and racial bias, poorly validated and improperly used, according to a recent report by NYPIRG, the New York Public Interest Research Group.

A review of NTE data in the public domain, conducted by NYPIRG's Blair Horner and Joe Sammons with the assistance of FairTest staff, uncovered "startlingly low" pass rates on the NTE for women and minorities (see box, page 7). Incredibly, the validity study to determine if New York should use the NTE was conducted by the Educational Testing Service (ETS), the company that produces the exam. Yet the state requires all teaching applicants to pass the test, and New York City can remove an existing teacher's license for failure on the NTE within five years even if classroom performance is satisfactory.

(continued on page seven)

CONTENTS

NAEP in Conference 3
Larry P. Case 4
WISC-R Questions 5
SATs & Blacks 9
Coaching Effective 9
Testing Children 10
Testing in States 10
Oops! 11
Publications/Events 12
Testing Problems 12
Order Form 13
FairTest Conference 15

Beyond Standardized Tests

(from page one)

AVERAGE 1987 ACT SCORES BY GENDER AND ETHNICITY

	MALES	FEMALES	TOTAL
Whites	20.4	18.9	19.9
Asian-Americans	20.3	19.3	19.8
Puerto Ricans & Other Hisp.	17.8	16.1	16.9
Native Americans/Alaskans	15.3	14.1	14.6
Mexican-Americans/Chicanos	16.2	14.8	15.4
Blacks	13.8	13.2	13.4
National Average	19.5	18.1	18.7

congressional committee: "We worry about people defining their worth and potential in terms of test scores. This is especially troubling because of race and gender differences in scores. We worry about the growing industry of test coaching schools feeding off of people's anxieties." MIT weighs SAT scores differently for men and women because research has shown that the test systematically underpredicts the future grades of females.

"We worry about people defining their worth and potential in terms of test scores. This is especially troubling because of race and gender differences in scores."

Favorable Reaction

FairTest found that applications increased at each of the schools after admissions reforms were adopted. A study conducted at Bowdoin College after submission of test scores was made optional also "suggests that the policy had the effect of expanding the applicant pool rather than simply substituting candidates with lower SAT scores for candidates with higher scores."

At Bowdoin, 98% of students responding strongly approved the new test policy, according to Admissions Director William Mason. The reaction was similar when Harvard Business School decided to stop considering Graduate Management Admissions Tests (GMAT) scores. "Our alumni, with their knowledge of what is important in a general manager, understood that the test scores were not helpful to us," concluded former Admissions Director John Lynch.

Based on the successful experiences of the seven schools, FairTest urged more colleges to independently evaluate their need for these tests. "If Harvard Business School can prosper when it refuses even to look at test scores, then other schools can certainly follow suit," concluded Executive Director John Weiss.

New Allies for Admissions Alternatives

The call to de-emphasize standardized tests recently won support from Robert H. Atwell, head of the American Council on Education, which represents 1,450 American colleges and universities. Speaking to a small West Coast conference whose participants included College Board President Donald Stewart, Atwell labelled the use of such tests "totally inappropriate" since they encourage schools to be "overly quantitative" in their admissions decisions. "Society should value the people who work hard, not just score well," Atwell concluded.

How to Order

Copies of *Beyond Standardized Tests: Admissions Alternatives That Work* are available from FairTest for $5.50 postpaid. Use the form on the inside back cover to place your order.

The National Center for Fair & Open Testing (FairTest)
P.O. Box 1272, Harvard Square Station
Cambridge, MA 02238; telephone (617) 864-4810.

Staff: *Mike Friedman*, Assistant; *Jim Gray*, Washington Liaison; *Noe Medina*, K-12 Consultant; *Mary Mengel*, Assistant; *Monty Neill*, Managing Director; *Bob Schaeffer*, Communications; *Sarah Stockwell*, Sex Bias Coordinator; *John Weiss*, Executive Director.

We would like to thank the Ford Foundation and the Rockefeller Family Fund for supporting our campaigns to ensure that the standardized tests administered to students and job applicants are fair, open and educationally sound.

FairTest Wins NAEP Reforms: More Problems Remain

In December, the U.S. Senate passed legislation to radically alter and expand the National Assessment of Educational Progress (NAEP), the federal government's major public school testing program started two decades ago. Though some safeguards advocated by FairTest are included in the Senate bill, its version still is flawed.

The NAEP Improvement Act would expand NAEP authorization from under $4 million annually to more than $85 million over a five year period. It would also allow up to 700,000 students to be tested annually in history, civics, geography, writing, science, reading and/or math.

"Whoever controls these powerful tests will control to a large measure what is taught and learned in American schools," George Madaus, Director of Boston College's Center for the Study of Testing, told *The New York Times*. For the past decade the Educational Testing Service has produced the NAEP tests under the supervision of a policy panel selected by ETS. In 1987, Secretary of Education William Bennett awarded ETS a $10 million contract to run NAEP until mid-1990.

Changes in the Legislation

Neither the Senate nor the House held any public hearings on the plan to expand NAEP, forcing FairTest and other civil rights and education groups to bring their concerns directly to key legislators. In November and December, FairTest staff met repeatedly with aides to key Senate and House legislators, particularly the bill's sponsor, Sen. Edward Kennedy. As a result of these meetings, the Senate legislation now includes the following provisions:

- *bias reduction techniques designed to help ensure that individual test items are not unfair to any ethnic, racial, gender or geographic group;*
- *the release of item analysis data documenting how different ethnic, gender and geographic groups perform on each item;*
- *stronger "Truth-in-Testing" provisions, forcing the release of more questions used on past tests;*
- *requirements that the policymaking board overseeing the NAEP contract include racial, ethnic and gender diversity;*
- *requirements that the NAEP policymaking board have paid staff independent of the contractor;*
- *requirements that the Secretary of Education provide for review by Congress a detailed plan of the program to expand NAEP before it is implemented.*

FairTest's Continuing Concerns

The proposal to expand and alter NAEP is now in House-Senate conference committee. It was included in the Senate version of H.R. 5, an omnibus education bill, but not in the House version. Thus, there remains some opportunity for further changes.

> ***"Whoever controls those powerful tests will control to a large measure what is taught and learned in American schools."***

FairTest believes three additional reforms are needed:

- *the new Act would, for the first time, allow state-by-state comparisons. Such comparisons previously were barred by Congress because of fears that competition between states to boost test scores would force teachers to alter their curricula to teach to the test, and that NAEP could dictate a national curriculum;*
- *the Act does not achieve its goal of making its policymaking board more independent of both the Department of Education and Educational Testing Service. Because two-thirds of the members of the current ETS-dominated Board will be automatically carried over to the new Board, its independence may be undermined;*
- *though Congress will now consider NAEP plans before they are implemented, the Act fails to provide for continuing public oversight of NAEP. Such oversight is necessary to provide an "early warning" of any potential harm the testing program could do to students, teachers and schools.*

FairTest will join other concerned organizations and individuals in urging the House and Senate conferees on H.R. 5 to consider these problems. If you are interested in supporting this effort, contact us.

FairTest Memoranda

FairTest's memoranda outlining in detail our concerns about the expanded NAEP, changes that have been incorporated as a result of our suggestions, and our recommendations for further changes is available for $4.

See also *The Nation's Report Card: Improving the Assessment of Student Achievement*, by the Alexander James Study Group with a Review by a Committee of the National Academy of Education, available from NAE, Harvard Graduate School of Education, 108 Longfellow Hall, Cambridge, MA 02138, for $9 a copy.

Larry P. Victory Under Assault

The landmark *Larry P.* decision, which bans the use of IQ tests in placing Black school children in "Educable Mentally Retarded" (EMR) classes in California, is under attack. Civil rights opponents have joined with proponents of standardized tests to use a "test case" to try to reverse this victory.

In *Larry P. v. Riles*, Federal Judge Robert Peckham agreed with parents who claimed that IQ tests are "culturally biased" and were being used unfairly to place large percentages of Black children in EMR classes. His ruling, first issued in 1979 and finalized in 1986, was twice upheld by the U.S. Court of Appeals.

Last May, Mary Amaya received a letter from her son's school claiming that they were unable to give him a battery of tests due to the Peckham decision. Amaya, claiming racial discrimination, brought her case to the newspapers. William B. Allen, a Reagan appointee to the U.S. Civil Rights Commission, took up her case and urged California to pass legislation allowing the tests to be given at a parent's request.

But Drs. Harold Dent and Gerald West of the Association of Black Psychologists explained that the Amaya case is being pushed by school psychologist organizations and reflects "the poor communication which the state has given to the various school districts" about the settlement. Peckham, explained Dent, stipulated that "if the state Department of Education wanted to request use of any given instrument, all they had to do was present data that the test was not biased. But because they cannot produce data to indicate that standardized tests of intelligence are not biased, they are trying to overturn the decision in some other fashion."

In the case of Mrs. Amaya's son, the school could use other, non-biased methods of assessment or the parents could pay for a private testing, said Dent and West. In any event, they agreed, public money should not be used to further discriminate under the guise of equal access.

Dr. Harold Dent, Dr. Asa G. Hilliard, III, Dr. Gerald West (left to right)

> ***"Any test that does not meet criteria of being fair, non-discriminatory and validated for the purpose of placing Black children in special education should be challenged in a local court."***

Dent and West believe the *Larry P.* ruling should be a model for other states. "Any test that does not meet criteria of being fair, non-discriminatory and validated for the purpose of placing Black children in special education should be challenged in a local court," explained Dent. The need for legal action is strong because "two thirds to three-quarters of the Blacks and Hispanics in special education programs are inappropriately placed." A separate suit in California, settled without trial in 1972, resulted in an agreement that the proportion of Hispanic children in special education would not exceed their proportion in the school district. They remain slightly over-represented in EMR, but very under-represented in "gifted" classes.

As for the challenges to the *Larry P.* decision, West observed, "It is very interesting how at this juncture the Office of Civil Rights is getting concerned about this instrument and over the years has not expressed a similar concern about the misplacement and labelling of Black youths. When we talk about these tests, we are talking about the institutionalization

(continued on next page)

Larry P. Case Under Assault

(from previous page)

of a diagnostic process that has no proven validity. Its sole justification appears to be economic and political expediency for the educational placement of youth whom California and the various states have not prioritized in terms of educating."

Concluded Dent, "The testing industry would like the general public to believe that standardized tests of intelligence are scientific and objective instruments, which in fact they are not." He cited three forms of bias: standardization (which is based on White middle-class norms), test administration (e.g., Blacks perform better when the administrator is Black), and biased items (see below for examples).

> ***"Two thirds to three-quarters of the Blacks and Hispanics in special education programs are inappropriately placed."***

For further background on the Larry P. case, see: *Negro Educational Review*, April-July 1987 (Box 2895, General Mail Center, Jacksonville, FL 32203), $4.

Examples from the WISC-R I.Q. Test

1) What are you supposed to do if you find someone's wallet or pocketbook in a store?
2 point response: Return it...Give it to the store owner (or a) policeman.
1 point response: Try to find the owner.
0 point response: Make believe you didn't see it. Keep it.

Is a poor child who keeps the money really less intelligent? Harold Dent, chair of the Testing Committee of the Association of Black Psychologists, argues that many minority children might not pick up the wallet or pocketbook for fear of being accused of stealing it. Such a response would earn the child no points.

2) What is the thing to do if you lose a ball that belongs to one of your friends?
2 point response: Give him/her one of mine...Buy another one if I can't find it.
1 point response: Try to find it
0 point response: Tell him you're sorry.

If you're poor, it often is not possible to buy another one and it is unlikely you have one to give. If the ball is really lost, looking for it is a waste of time; and if it is not really lost but can be found, wouldn't looking for it be more "intelligent" than buying a new one? Apologizing might be the best response for a poor child, but such an answer will yield a lower "intelligence" score.

3) How tall is the average American man?
Any answer from 5'7" to 5'11".

Children often answer the height of their fathers. Asians and Hispanics are often shorter (and Latin *Americans* are a majority of residents of the Americas).

4) Who discovered America?
Acceptable responses: Columbus, Leif Ericson, Vikings (Norsemen), Amerigo Vespucci.

A response of "Indians" prompts the tester to say, "Yes, the Indians were already there. But who sailed across the ocean and discovered America?" The claims of West Africans or Polynesians or Irish that their sailors reached the shores of "America" long before Columbus are not included as correct answers, but all of these responses are surely more accurate than "Vespucci," who did not *discover* the continents that now bear his name.

In this section of the test ("Information"), two other white men were specifically mentioned, but no woman or person of color is directly referenced *anywhere* in this section. The message is that an informed child knows about white men, but does not need to know about women or people of color.

5) What is the thing to do when you cut your finger?
2 point response: Put a bandaid on it.
1 point response: Go to the doctor (or) hospital.
0 point response: Cry, bleed or suck on it.

Minority children usually perform poorly on this item. A few years ago, Johns Hopkins University researcher Janet Hardy asked 4,000 inner-city youths why they answered WISC-R questions the way they did. She found that on this item, many of these kids answered "go to a hospital" because they thought that "cut" meant a big cut. Or they thought it was a small cut, and since they didn't have any bandaids at home, they answered "sucked on it"—and received no points.

Albany, New York November 20, 1987

TIMES UNION

Teacher test called biased

Minorities score 'abysmally low'

By Tim Spofford
Staff writer

The examination that prospective teachers must pass to get jobs in New York state is biased against women and minorities, the New York Public Interest Group charged Thursday

NYPIRG argued that test scores for minorities on the National Teachers Examination are "abysmally low" and that women are twice as likely as men to fail the section on general knowledge. "The exam adversely impacts on women and minorities, and we think there is some inherent bias," said Blair Horner of NYPIRG

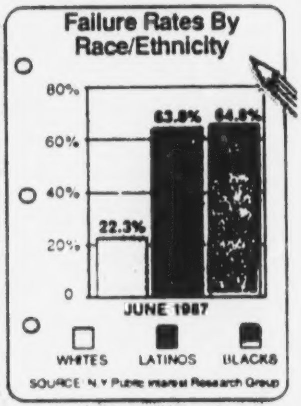
However, the state Education Department challenged the findings in NYPIRG's report, titled "The Test That Fails"

"There is no evidence that I've seen or read about to indicate the test is biased against minorities or women," said Charles Mackey, supervisor of teacher education in the state's Division of Teacher Certification

Used by 17 states to screen teachers for jobs and certification, the National Teachers Examination has sparked controversy and several lawsuits. Some critics of the exam argue it is too easy and permits unqualified candidates to slip into teaching. Others charge it discriminates against blacks and Hispanics and is thinning the ranks of minority teachers.

The National Teachers Examination is developed by the Educational Testing Service of Princeton, N.J. — the same firm that develops the Scholastic Aptitude Test and many other standardized tests.

Failure Rates By Race/Ethnicity



SOURCE: N.Y. Public Interest Research Group

Times Union Graphic

The teachers exam includes tests one each in communication skills, general knowledge and professional knowledge. The communications test evaluates such skills as reading and writing. The test of general knowledge covers subjects such as literature, math and science. The third test is on teaching methods

Since 1985, teachers entering the profession in New York state have had to pass all three tests to get jobs and earn certification. But many minorities and women have found it tough to clear the hurdle.

According to figures compiled by NYPIRG, the passing rate for blacks who took the exam in New York in June 1987 was nearly half that of the whites. Just 36.1 percent of the blacks passed the test of communications

See **SURVEY** / A-10

SURVEY

Continued from A-1

skills, 35.2 percent passed the test of general knowledge and 56.5 percent passed the section on professional knowledge.

Whites who took the exam had a passing rate of 83.3 percent on the test of communication skills, 77.7 percent on the general knowledge section and 91.3 percent on the test of professional knowledge.

Women who took the test of general knowledge in June 1987 had a passing rate of 66.7 percent, while men had a passing rate of 83.2 percent, NYPIRG reported.

Horner said these results show the exam is biased, but since he cannot get copies of the tests, NYPIRG cannot state where the bias lies. "It's difficult for us to say definitely what the problem is with the exam," he said.

But Mackey argued the exam is fair to all test takers. He said the lower scores for minorities may be caused by the fact that cities such as New York and Buffalo recruit minority teachers from the South, Puerto Rico and the Dominican Republic — regions with lower educational standards.

When these teachers take the exam in New York, they lower the average scores for minorities, Mackey said

Horner disagreed. "The problem is you find that failure rate (among minorities) in other states, and they're not all recruiting out of the country"

Women may have lower scores on the test of general knowledge because it has a challenging math and science section, Mackey said, adding that women tend to take fewer courses than males in those two subjects.

But Horner countered that girls outperformed boys on two Regents math tests studied recently by the state Education Department. The problem lies not in women test takers, but in the Educational Testing Service, he added.

Horner criticized the department for hiring the service to validate the teachers examination for use in New York. Because the test was written and sold by the firm, it invited a "conflict of interest" by handling the study, Horner added.

Mackey defended the state's validation study, arguing that it relied chiefly on the judgments of teachers and professors from New York. The company merely compiled these observations to write the final report, he said.

ETS declined to comment on NYPIRG's report. "We really can't comment on the report as we don't have a copy," said spokeswoman Jana Cardoza.

NYPIRG's report called for a special Truth-in-Testing law to make testing firms return exams with answers to all test takers in professional fields such as teaching, medicine, law, architecture and others with certification that depends on test results. NYPIRG called for a second law creating a state commission to study all professional tests for cultural or gender bias.

NYPIRG also called for an independent study to examine the validity of using the examination to certify New York teachers.

Connie Eno, president of the National Education Association of New York, urged the state Education Department to hire an independent firm to review the exam. "We have some grave questions about the test — that it may be biased against women and minorities," she said.

Eno added that questions on the professional knowledge portion of the test are often unclear, and more than one answer is often possible. "In many cases, there are about two or possibly three correct answers," she said. "The answers are arbitrary."

Peter Boespflug, spokesman for New York State United Teachers, said he has not seen NYPIRG's report. "But if there are inequities, they certainly should be rectified," he said.

Teacher Test Flunks

(from page one)

Like other researchers, NYPIRG found no evidence that the NTE separates competent teaching applicants from incompetent ones. Indeed the test's manufacturer does not claim predictive validity for its product.

To counter the NTE's flaws, NYPIRG proposed a "Code of Fair Teacher Testing Practices" including implementation of Truth-in-Testing safeguards against ambiguous, flawed, or inaccurately scored questions, and the appointment of a study committee, similar to that established for college and university admissions tests (*FairTest Examiner*, Vol. 1, Nos. 1, 2 & 3), to investigate charges of bias. NYPIRG also called for independent studies of the NTE's validity and use.

Shortly after release of NYPIRG's report, the New York State Education Department announced plans to conduct a major study of occupational licensing tests under its control, including the NTE. The $250,000 research project will look at test-takers rights, bias protections and the relevance of the exams to on-the-job performance. A contract to conduct the research will be awarded in June, 1988. A letter-of-intent from prospective bidders is due February 1. The final proposal deadline is April 1. Contact Daniel Szetela at (518) 474-3863 for copies of the Request for Proposals.

National Implications

Currently 17 states use the NTE for teacher certification and licensing. The impact of the test is similar across the country. As ETS President Greg Anrig has admitted, if states continue requiring passage of the NTE for teacher certification, "ETS research reports conclude that, by the year 2000....the percentage of minorities in the teaching force could be cut almost in half," at a time when the number of minority students is growing rapidly.

Reacting to charges of bias in the exam, Rhode Island recently set up a commission to examine NTE items for race and gender bias. In Alabama, a landmark consent decree settling a racial bias lawsuit led to the complete overhaul of that state's teacher test, which is made by National Evaluation Systems (details in *FairTest Examiner*, Vol. 1, Nos. 1 & 3).

In Georgia, teachers have launched a suit to cancel results of a test administered to all currently certified teachers in the state. Of the 327 educators who did not pass, 244 are black. "Are these tests valid, job-related and non-discriminatory?" asked Georgia Association of Educators President Faye Fox. "Based on the available evidence, we believe the answer is no." The case is expected to go to trial next summer.

> ***"Are these tests valid, job-related and non-discriminatory? Based on the available evidence...the answer is no."***

PERCENT PASSING NTE GENERAL KNOWLEDGE TEST
(New York State,
June, 1987, Administration)

	PASSING RATE
Black	35.2%
Asian	57.7
Puerto Rican	31.6
Other Hispanic	43.6
White	77.7
Male	83.2
Female	66.7

Academics Criticize Teacher Tests

Several recent articles in leading educational publications provide more ammunition for critics of the NTE and other teacher tests. Writing in the September, 1987, issue of *Phi Delta Kappan*, George F. Madaus and Diana Pullin of Boston College found, "The generic [teacher licensing] tests that we have examined to date convince us that test developers have failed miserably in their attempts to define this elusive domain of common knowledge and to write items that measure it."

Even stronger is Madaus' conclusion writing with Boston College colleagues Walter Haney and Amelia Kreitzer, in *Review of Research in Education* (Winter, 1987). Their article, "Charms Talismanic: Testing Teachers for the Improvement of American Education," concludes, "(Our) results indicate clearly that current teacher tests, and the manner in which cut-scores are being set on them, are differentiating among candidates far more strongly on the basis of race than they are on the basis of teacher quality...This suggests to us a modest proposal: do not use such tests and cut scores until it is clearly proven that these tests select among teacher candidates more on the basis of some independent measure of teacher quality than on the basis of race."

The Test That Fails is available from NYPIRG Publications, 9 Murray Street, 3rd Floor, New York, NY 10007 for $5.00.

The Boston Globe

SATURDAY, NOVEMBER 7, 1987

Scratching the SATs

High school students in Massachusetts and across the country will be participating in the autumnal rite of the Scholastic Aptitude Tests this morning – just as many colleges are recognizing that the tests are about as important as football rallies and other traditions of autumn.

The SATs are supposed to help colleges judge the comparative standing of applicants from a wide variety of schools. What has become increasingly clear is that the tests do not provide a good means of judging the comparative standing of applicants from a wide variety of social and economic backgrounds.

The social-economic bias of the tests has been demonstrated by studies that show they accurately reflect the income and educational attainment of an applicant's parents. This calls into question any judgment that can be made about the scholastic aptitude of the student himself on the basis of test scores.

The problem has been compounded in recent years as it has become clear that a student's scores can be increased dramatically by tutoring. Few public high schools provide SAT tutoring as part of their regular class program. Private tutoring programs cost up to $500 – which puts them out of the reach of students from low-income families, the very students who most need the tutoring. This cost comes on top of the cost of taking the test ($12) and of sending it on to college admissions offices.

At Massachusetts Institute of Technology, the admissions director worries "about the growing industry of test-coaching schools feeding off people's anxieties."

Recognizing these problems, a number of the nation's top-ranking colleges – including Bates, Bowdoin, Middlebury and Union colleges, and the Harvard Business School – have dropped or deemphasized the SATs and other standardized admissions tests. According to a study by the National Center for Fair & Open Testing – FairTest – "all of these schools report that the reaction to these changes is overwhelmingly positive."

At Bowdoin, where submission of test scores has been optional for a number of years, the admissions office reports the applicant pool was expanded by "more highly motivated but low-scoring students."

Harvard Business School has gone the furthest and now does not even look at Graduate Management Admissions Test scores, even if an applicant submits them. As FairTest notes, if the B-School "can prosper when it refuses even to look at test scores, then other schools can certainly follow suit."

That is a proposition that makes considerable sense. Certainly, if the tests are going to be required of applicants, then the colleges should be willing to pay for them – and the College Board, which administers the tests, should provide financial assistance for low-income students to take private tutoring courses.

Drug Boosts SAT Scores

A drug used to control high blood pressure has dramatically improved SAT scores of students who suffer from exceptionally high levels of test anxiety, according to Brandeis University researcher Dr. Harris Faigel. After the students took Propranolol, a beta-blocker used to lower high blood pressure and lower heart rate, their scores improved by a mean of 50 points on the verbal and 70 points on the math. Typically, "students who retake the test without special preparations increase their verbal scores by 18 points and math scores by 20 points," Robert Cameron, Director of Research for the College Board, told *The New York Times* of October 22, 1987. Dr. Faigel's study will be published in the Spring 1987 issue of *Pediatrics*.

SATs Underpredict Blacks' College Success

"SAT scores consistently predict failure when in fact kids succeed." So says Timothy L. Walter, assistant professor of Physical Education at the University of Michigan and co-author of a new study on Black athletes and college performance.

Tracy Graham, a freshman at Iowa State University, lost her eligibility to play volleyball because she took the ACT in July 1986, a date not approved by the NCAA for taking the test. Graham lost one appeal to the NCAA, but may regain eligibility when the NCAA reviews her case in January. The NCAA regularly restores eligibility to male star athletes who commit far more serious infractions of NCAA rules.

To compete in interscholastic athletics under current NCAA rules, incoming freshmen must have a 2.0 (out of 4.0) grade point average (GPA) in a high school core curriculum, and score at least 700 on the SAT or 15 on the ACT. Many minority athletes meet the grade requirements but are declared ineligible due to low test scores.

But Walter found SAT scores were uniformly poor at making accurate predictions about college grades for Black football players who entered the University of Michigan between 1977 and 1983. While high school GPA correctly predicted success in 84% of the cases, test scores (SAT and ACT) accurately predicted success only 30% of the time. A combined test score/GPA index, which is now used by the NCAA, was correct less than half the time.

Finally, the study found that 86% of Black student athletes who would have been denied entrance to Michigan because of low test scores did in fact succeed in their freshman coursework, and a large percentage of those went on to graduate.

Professor Walter concluded, "By overpredicting failure, (test scores) keep a whole group of athletes out who deserve to be there. The SAT just doesn't do the job as an indicator."

Though the study did not look at the larger issue, it raises the question of how many minority students who are not athletes are denied access to colleges and universities by biased test scores.

(Source: *Chronicle of Higher Education*, November 4, 1987.)

Coaching Effectiveness: The Evidence Mounts

The College Board and Educational Testing Service continue to deny that short-term preparation can significantly boost SAT scores. But an increasing body of evidence demonstrates that testmakers' claims are little more than self-serving promotion. Among the new evidence:

The **National Urban League** reports that its SAT coaching program has boosted the scores of 29 students in Harrisburg, PA by an average of 70 to 100 points. A letter from the **University of Pennsylvania's Upward Bound Program** asserts that their students enrolled in the Princeton Review SAT preparation course showed an average improvement of 160 points over their prior scores.

Father Richard Curran, a Boston University doctoral student who is examining commercial coaching software packages, reports that scores increased an average of 57 points on the SAT verbal and 40 points on the math for the 46 juniors in his pilot study. Curran is in the process of analyzing further data from 204 parochial school students.

Copies of **John Zuman's** Harvard Graduate School of Education thesis, "The Effectiveness of Special Preparation for the SAT: An Evaluation of a Commercial Coaching School," are available from University Microfilms, Ann Arbor, MI. Zuman, who found a 110-point average increase in SAT scores after test coaching by the Princeton Review program (see *FairTest Examiner*, Fall 1987), will be presenting the results of his study at the 1988 American Educational Research Association conference.

TESTING CHILDREN

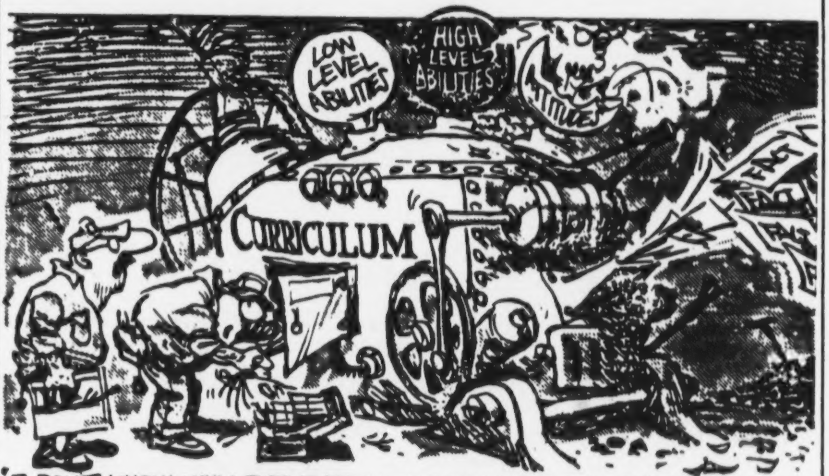
Tests Narrow Curriculum

The Southern Education Fund (SEF) has released *School and Competency Testing Programs: Perceptions and Effects on the Black Students in Louisiana and North Carolina.* The report charged competency tests with narrowing curriculum and causing teaching to the tests, and argued that remedial programs that pull students from regular classes may cause students to fall further behind. In both states, Black students are still underrepresented in gifted and college prep programs. The report suggests that the focus on basic skills and testing has diverted effort from teaching the skills required by colleges. Further, the testing programs may be driving some Blacks out of school.

The study concluded that minority students and students with poorly educated parents fared less well on the tests. Of N.C. high school juniors taking the test from 1978-1983 whose parents studied beyond high school, 98% passed the reading and 97% passed the math tests, while for those whose parents' educational level was eighth grade or less, the respective figures were 71% and 70%. From 1978 to 1985, the first-time pass rate of Whites was 97% for reading and 95% for math, while the respective scores for Blacks were 82% and 77%.

Among the recommendations in the report: an end to test-score based ability grouping, a broadening of curriculum, more Black teachers, school program flexibility for diverse needs, and more college involvement in helping high school students.

(Order from SEF, 340 W. Peachtree St., N.W., Suite 250, Atlanta, GA 30308; $8.00.)



"I DON'T KNOW WHY, BUT EVERY TIME WE WIRE THE CURRICULUM STANDARDS TO THE TESTS, WE BLOW THOSE TWO BULBS AND A WHOLE MESS OF TRIVIA FLYS OUT ...."

All States Are Above Average

"[No] state is below average at elementary level on any of the six major nationally normed, commercially available tests!" according to a new study by the West Virginia-based Friends of Education. This absurdity is possible because norming is usually done every 8-10 years, and over time the average score rises as teachers learn to teach to the test.

Not only does this call into question the gains in scores that Education Secretary Bennett credits to the "back to basics" emphasis of the Reagan administration, it also calls into question the testing process in general. Is it preferable to teach testing tricks or reading, math and thinking skills?

(Order from Friends of Education, Box 358, Daniels, WV 25832-0358; $5.)

SAT Gender Bias

Preliminary research by the University of Delaware confirms a FairTest study concluding that the SAT is biased against women. "Women, on the average, out-perform men in both high school and college, yet they score lower on the SAT," said researcher Michael Middaugh. While several university officials defended use of the test, the student paper referred to the "S(exist) A. T." and condemned its continued use. Copies of the FairTest report, *Sex Bias in College Admissions Tests: Why Women Lose Out*, are available for $7.95. Use the order form inside the back cover.

TESTING IN THE STATES

Connecticut has developed an unusual system for apportioning a part of state education funds: The lower a community's scores, the more money it will receive. While the new formula avoids the common tendency to punish low-scoring districts for their residents' poverty, it may inadvertently encourage the use of tests which exaggerate the gaps between rich and poor.

The new **Massachusetts** Testing Program included among its regulations the following model Truth-in-Testing provision: "11. Each public school district, following receipt of information on individual student performance under the Basic Skills program, shall report to such students and their parents, complete information on their performance. This information shall include the opportunity to review the test instruments and the corrected responses of the student on these instruments."

"There is a relation between a school's wealth or poverty and (test) results."

The fate of **Bronx, N.Y.**, Junior High School 123, which outgoing Schools Chancellor Nathan Quinones had ordered closed due to poor test scores, remains in doubt. In September, Judge Bertram Katz ordered the school kept open, saying there was "serious doubt" about making administrative decisions based on test scores and therefore Quinones' decision was not rational. In December, a N.Y. Appeals Court overturned Katz in a split decision. The school board has not decided whether to close the school in mid-year and the appellate decision will be appealed by the Community School District. Parents groups, who acknowledge many problems at the school, have led the charge to keep it open. Said one parent, "The school was bad, but it was not one of the worst...If they were going to close our school, well, they should close the whole system down." (*NY Times*, 11/18/87, 12/18/87)

In **Pennsylvania**, controversy has erupted over the meaning of the recently-released scores for the Test for Essential Learning and Literacy Skills (TELLS), given statewide to students in grades 3, 5 and 8. Secretary of Education Thomas K. Gilhool argued that the tests show "that poverty does not have to be an obstacle to high academic achievement." His evidence was 47 schools with a high rate of poverty and a high rate of student mastery of both English and math; and a large number of schools, some with high poverty rates, showing 10 percent or more improvement.

However, the Pennsylvania League of Urban Schools (PLUS) released an analysis of the schools indicating "there is a relation between a school's wealth or poverty and TELLS results." The state figures did not correlate test performance with per-pupil expenditures. PLUS has argued that improving education requires increasing the money available to low-income districts. PLUS did succeed in persuading Gilhool not to release a ranking of the school districts, explaining that such a comparison was not statistically valid and would take emphasis away from more significant aspects of the data.

Summaries of interviews with teachers by the state Dept. of Education indicates that some of the rising scores may be attributable to "teaching to the test," and "alignment" of the curriculum, although other factors such as increased parental involvement were consistently mentioned as well.

The Atlantic Center for Research in Education (ACRE), the group which successfully led the fight to stop mandatory standardized tests in grades 1 and 2 in **North Carolina**, reported, "There are at least three state agencies who want to reinstate achievement testing in first and second grades." ACRE is gearing up to preserve their victory. (For further information about the successful campaign in N.C., contact: ACRE, 604 W. Chapel Hill Rd., Durham, NC 27701, (919) 688-6464.)

WORK FOR FAIRTEST

ADMINISTRATIVE ASSISTANT

Responsible for: office support, including light bookkeeping, mailing, correspondence, filing, data entry and maintenance, managing publication orders and subscriptions. **Assist with:** producing FT publications, logistics for events, and research projects.

Requirements: Office experience, writing skills, type 50wpm and computer experience. Candidate should feel comfortable working in an informal setting for small organization.

Salary: $15,000 to $18,000/year. Health benefits and three weeks vacation.

Summer internships and work study positions also available.

Equal Opportunity Employer

Oops!

New York Bar Exam

Can one question make a difference? As a result of Steven Shapiro's challenge of one ambiguous item on the July 1987 New York Bar Exam, 26 law school graduates who had failed the exam have been notified that they passed the test and 33 others will be able to appeal their test results. Said one of the 26, "A good portion of the bar exam is nothing more than a memory test and a crapshoot."

This mistake marks the third consecutive year that scores in New York have had to be changed because of errors. Last year, the fortunes of 22 applicants improved after a faulty question was discovered on the exam. In July 1985, 111 test-takers had failing grades reversed because of a computer error in scoring their tests, and 542 others had to re-take the test when their multistate answer sheets were lost.

N.Y. City Police Test

In November, a New York Judge upheld challenges to 21 of 38 challenged items on the 1985 test required by New York City for promotion to police Captain. Judge Arthur Blyn ordered that the scoring key be changed on 10 of 19 challenged questions on the administrative portion of the exam and 11 of 19 questions on the technical knowledge section. The ruling will likely be appealed, but if it stands, those who passed the exam with the revised key will be able to take the oral part of the Captain's test and will then be inserted in the hiring list in order of their final exam scores.

Publications to Note

The Negro Educational Review of April-July 1987 is a special issue, edited by Asa G. Hilliard, III, *Testing African American Students*. Among many excellent articles is "It's Time to Examine the Examiners" by FairTest Executive Director John Weiss. Order from Box 2895, General Mail Center, Jacksonville, FL 32203; $4.00.

The Fall 1987 issue of *Educational Measurement: Issues and Practices* includes ETS President Greg Anrig's reaction to the debate on the Golden Rule bias reduction procedure printed in that publication's Summer 1987 issue. The Winter 1988 edition of *EM* includes a response from Golden Rule's Chief Executive, Pat Rooney. Copies of the journal are available for $5.00 (subscription $12) from NCME, 1230 17th Street, N.W., Washington, DC 20036. The Anrig and Rooney articles will be reprinted in the next issue of FairTest's **Update on Testing Reform Initiatives**.

What is the Appropriate Role of Testing in the Teaching Profession?, proceedings from the NEA, AACTE, ETS and FairTest joint conference, has been printed by the National Education Association. FairTest Executive Director John Weiss's paper, "Testing Teachers: Strategies for Damage Control," is included, as well as other featured papers by ETS President Greg Anrig, AACTE head Norene Daly, and Sharon Robinson of the NEA. Copies are $8.95 from NEA, Box 509, West Haven, CT 06516.

You may be interested in ordering *Rethinking Schools* (Box 93371, Milwaukee, WI, 53202; bimonthly, $5/yr), an excellent newspaper covering not only Milwaukee but many important education issues for the nation as a whole.

Upcoming Events

FairTest's Executive Director John Weiss will be a panelist at a session of the **National School Boards Association** convention in New Orleans on March 28th, 1988 from 1-2 p.m. The other panelists invited are Wilmer F. Cody of the Council of Chief State School Officers, J. Douglas Conner, executive director of the American Association of Collegiate Registrars and Admissions Officers, and Gregory Anrig, president of ETS. The topic will be the status and uses of standardized student testing.

The American Educational Research Association and National Council for Measurement in Education will hold their 1988 annual meeting in New Orleans, LA, April 5-9, 1988. Address inquiries to the AERA Meeting Coordinator, 1230 17th St., N.W., Washington, DC 20036.

The National Association for Bilingual Education's Seventeenth Annual International Bilingual/Bicultural Education Conference is titled "Bilingual Education: Excellence Touches the Future." The conference will be held in Houston, TX, April 25-29, 1988. For more information contact: Dr. Sylvia Pena, NABE Program Co-Chair, C&I College of Education, University of Houston, Houston, TX 77004.

Problems with Test Administration

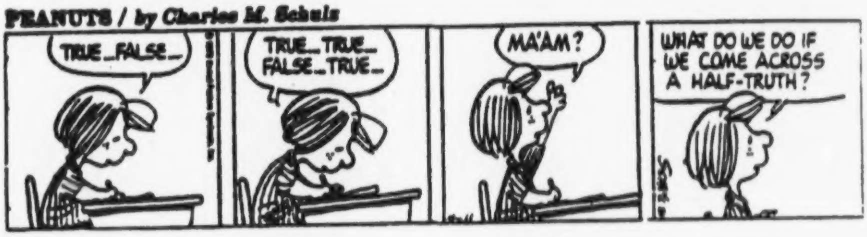
In an analysis of 22 controlled studies, Vanderbilt University researchers Douglas and Lynn Fuchs found that familiarity with the test examiner significantly raised test performance. This was particularly true when "subjects were of low socio-economic status (SES) [and] were tested on comparatively difficult tests." High SES children, however, were unaffected by examiner familiarity. The examiner effect could account for 8 points on a standardized IQ test with a mean of 100 and a standard deviation of 15, or about half the average score difference between low and high SES children. (Fuchs and Fuchs, *Review of Educational Research*, Summer 1986.)

In a subsequent meta-analysis, the Fuchs found that the effect of examiner familiarity on minority-group children is significant: unfamiliar examiners harm the performance of Black or Hispanic children (*Chronicle of Higher Education*, 11/18/87:A15). Both the *Standards for Educational and Psychological Tests* and most test makers discourage examiner familiarity and thus effectively encourage systematic bias against low SES and minority students.

More Problems

A "study of teachers' group testing practices in eight kindergartens" by Kenneth Wodtke, *et al*, of U. Wisconsin-Milwaukee, "revealed wide variation in testing conditions [and] a high frequency of discrepancies from the standardized administration procedures." The practices "were so nonstandardized as to render their scores...possibly unreliable." The scores, however, were used for pupil placement, "resulting in potentially injurious educational consequences."

PEANUTS / by Charles M. Schulz


TRUE...FALSE...
TRUE...TRUE...
FALSE...TRUE...
MA'AM?
WHAT DO WE DO IF WE COME ACROSS A HALF-TRUTH?

The authors concluded that if such non-standardized conditions are prevalent, it "would represent a serious case of malpractice in early school testing." They expressed concerns about "the *indiscriminate* group testing of kindergarten children," wondered if the standardization expected by test-producers can be attained, and argued that if it cannot, the continued use of the tests would "perpetuate a form of pseudoscience... harmful to children's educational development."

Racial Tracking in Public Schools

As the proportion of Black teachers increases, the proportion of Black students suspended, expelled or placed in special education classes declines, according to a recent analysis of 173 urban school districts. In general, Blacks are more than twice as likely as Whites to be in special education and half as likely to be in classes for the gifted. The use of standardized tests helps ensure the continued misclassification of Blacks. These results confirm previous studies, including the research done for the *Larry P.* case (see story on page 4).

The researchers, Kenneth Meier, Robert England and Joseph Stewart, will soon publish *Race, Class and Education: The Politics of Second Generation Discrimination*. This study also adds to the mounting evidence on the importance of increasing the number of Black teachers at a time when those numbers are declining, largely due to low scores on standardized exams for teachers (see *Teacher Test Flunks*, p.1).

Another High School Test?

Under the sponsorship of Sen. Claiborne Pell, the U.S. Senate, in its version of omnibus education bill H.R. 5, has voted authorization for the Department of Education to spend $2 million to develop "The Optional Test of Academic Excellence." High school juniors would have the option of taking the test, and those who passed would receive a certificate from the Secretary of Education. It is currently unclear whether students could submit their scores to college admissions officials.

"Any test written by the Secretary of Education could turn out to be an instrument to bludgeon school officials, not a way of recognizing students," Richard Kruse of the National Association of Secondary School Principals told *The New York Times*.

As with the NAEP Improvement Act (see article, p. 3), the "Excellence Test" is in conference committee. FairTest urges that the committee delete this provision from H.R. 5.

Sunday, November 29, 1987 New York Daily News Circulation 1,750,000

SATs harm: (A) women & (B) minorities?

By CAUSEWELL VAUGHAN
Daily News Staff Writer

The Scholastic Aptitude Test that 200,000 area high school students will take next weekend often yields scores that are biased against minority-group members and females and should be abolished, according to a growing number of critics.

The test, first used in 1926 to measure intelligence, now is employed by many of the nation's colleges as a kind of entrance exam. In New York, millions of scholarship dollars are assigned according to test results.

But critics—among them several colleges that have dropped the test—claim it is skewed in favor of white, middle-class males and should be thrown out.

"It is time to overhaul the Scholastic Aptitude Test to eliminate its biases," said Bob Schaeffer. "If that can't be done, the exam should be discarded."

Schaeffer is the public education consultant for the National Center for Fair & Open Testing (FairTest), a nonprofit organization based in Cambridge, Mass. It has been the leading critic of the SAT.

FairTest claims that females and minority-group members who take the multiple-choice SAT not only don't score as well as white males, but lose out on millions in scholarship money geared to test results. The bias, FairTest says, is in the questions themselves.

In a study, "Sex Bias in College Admissions: Why Women Lose Out," FairTest deemed the questions biased because of their "rushed nature and multiple choice format": they feature more males than females, more emphasis is placed on subjects that favor males, and they focus mostly on male interests and activities.

Minority group females taking the test face double jeopardy, the study said, because many questions require students to be familiar with the language and activities of the upper-middle class.

1.5 million students take test

Nationwide, close to 1.5 million high school students take the test annually. It comprises two parts, verbal and math. Students' scores can range between 200 and 800 for each part, making 1600 a perfect SAT score.

Last year, females averaged 61 points less than males, scoring an average 50 points less on the math and 11 less on the verbal. According to FairTest, the gap between the sexes has been widening for 15 years.

The College Board, which administers the SAT, says it contains no bias. Speaking for the organization, Ann Grosso said, "Statements about bias in the test are simply inaccurate." The negative statistics, she said, "reflect a difference in educational experience."

Grosso said 88% of the nation's colleges use the SAT as an indicator of how students will do in their first year of college. Last year, the number of colleges using the test totaled 1,583, an increase of 400 colleges since 1977.

Four schools have dropped them

In recent years, however, four small colleges—Middlebury in Vermont, Bowdoin College and Bates College in Maine and Union College in upstate Schenectady—have dropped SAT scores as an admission requirement. A recent move for similar action at Brown University was barely defeated. The consensus is the SAT didn't work for the student or the institution.

Bates College dropped the test as of 1984. According to Director of Admissions Wylie Mitchell, there were "ethical concerns."

"The faculty was concerned about how the test related to students' family backgrounds," Mitchell said, "particularly poor kids whose scores reflect the income factor as opposed to the academic factor."

At Union College, next fall's freshman class will be the first admitted without the SAT as a requirement. Speaking for the college, Peter Blankman said, "We don't think the test can escape a cultural bias."

Officials at these colleges and others simply believe students' grade-point average should be weighed more heavily in admissions decisions, as opposed to focusing on the SAT.

Grade-point 'more important'

Dr. Silvia Johnson, a professor in the school of education at predominantly black Howard University, said, "Grade-point average is more important than the SAT. ... A lot more attention should be given to try to develop items that tap fairly into the ability of all kids."

Educators who don't have high regard for the test also point to the fact that students' scores can be distorted by coaching classes costing up to $500.

Choice samples

The following is a sample of SAT questions that FairTest claims are biased. Students are asked to choose the pairing most similar to the heading. The answers are below.

1. RUNNER: MARATHON
- (A) envoy: embassy
- (B) martyr: massacre
- (C) oarsman: regatta
- (D) referee: tournament
- (E) horse: stable

2. RACQUET: TENNIS
- (A) springboard: diver
- (B) horse: polo
- (C) glove: boxing
- (D) club: golf
- (E) gun: hunting

3. HEIRLOOM: INHERITANCE
- (A) payment: currency
- (B) belongings: receipt
- (C) land: construction
- (D) legacy: bill
- (E) booty: plunder

4. MELODEON: ORGANIST
- (A) reveille: bugler
- (B) solo: accompanist
- (C) crescendo: pianist
- (D) anthem: choirmaster
- (E) kettledrum: tympanist

5. BRIDLE: HORSE
- (A) bone: dog
- (B) olive branch: dove
- (C) valor: soldier
- (D) precept: conduct
- (E) devotion: duty

ANSWERS [illegible]

FairTest says close to 300 colleges have either absolute cutoff scores or specify a cutoff score as a part of admissions requirements. As long as the SAT is used in this manner, he says, women and minority-group members will lose out in education and in life.

New York State Merit Scholarships for about $40 million annually go to students with high SAT scores. The organization said that of 1,000 Empire State Scholarships of Excellence awarded in 1986-87, 720 went to males and 280 to females. Of the 109,266 students who applied for New York State Regents Scholarships, 47% were male and 53% female, but males won 57% of the money.

Phyllis Rosser, who compiled the FairTest study, said the statistics are damning for females and members of minority groups. "This means real dollar loss for them in later life," she said, "as they receive lower-level jobs, earn less money and have fewer leadership opportunities, and test discrimination can also create a lifelong loss of self-confidence that cannot be calculated in dollars."

Standardized Testing Reform Sourcebook 1988

Completely revised edition of FairTest's popular looseleaf notebook resource for advocates, lawyers and educators.. Includes key legislation, court decisions, reports and academic articles covering the Golden Rule Principle, Truth-in-Testing and Item Analysis. An essential reference for anyone working on testing reform.

$25.00 Postpaid
(approx. 200 pages)

use order form on page 15

Subscribe to the *FairTest Examiner*

The *FairTest Examiner* will keep you up-to-date on the growing movement for fair & open testing. You will be able to do more than just read about FairTest campaigns--you can help make them successful!

_____ *FairTest Examiner*, published quarterly...Individuals $15/year; Institutions $25

_____ *Standardized Testing Reform Sourcebook, 1988* (200 pages)...$25

_____ *Sex Bias in College Admissions Tests: Why Women Lose Out*,
by Phyllis Rosser and FairTest staff, 2nd edition (40 pages)...$7.95

_____ *The Reign of ETS*, by Allan Nairn and Ralph Nader (550 pages)...$30

_____ *Beyond Standardized Tests: Admissions Alternatives that Work*,
by Amy Allina with FairTest staff, (15 pages)...$5.50

Updates, issued bi-annually (spring & fall). $7/year/topic:

_____ *Testing Reform Initiatives*
_____ *K-12th Grade Testing*
_____ *University Admissions Tests*
_____ *Teacher Testing*
_____ *Employment Testing*
_____ *Sex Bias in Testing*

order total _____

tax-deductible contribution _____

TOTAL ENCLOSED _______
(enclose check to FairTest)

_____ **Back issues of *FairTest Examiner*** $4 each: Vol___, No.___ (all 3 for $12)

_____ **Moving.** I have moved/will be moving. My new address is (please enclose old label or write old address.):

Name ______________________________ Title ______________________________

Organization ____

Address ____ Zip __________________

Make checks payable to FairTest and mail to: FairTest, P.O. Box 1272, Harvard Square Station, Cambridge, MA 02238

FAIRTEST CONFERENCE

Breaking Down the Barriers

June 17 & 18, 1988

Washington D.C.

Friday, June 17, 4-7 p.m.:
Capitol Hill Reception
Keynote Speaker:
Representative Patricia Schroeder

Saturday, June 18, 9-5

Speakers and workshops will cover:
Teacher Testing,
Legislative and Legal Strategies,
University Admissions Reform,
Gender and Race Bias,
Testing Bilingual Students,
K-12 Initiatives,
Coaching Controversy,
"Intelligence" Tests,
and more.

Reserve these dates!
Further information will be in the next *FairTest Examiner*

FairTest
National Center for Fair & Open Testing
P.O. Box 1272
Harvard Square Station
Cambridge, MA 02238
(617) 864-4810

Bulk Rate
U.S. Postage
PAID
Cambridge, MA
Permit No. 53278

Complimentary Copy.
Please Subscribe today.
See page 15.

Address Correction Requested

FairTest

EXAMINER

The National Center for Fair & Open Testing Volume 2, Number 2 Spring 1988

John Weiss to step down as FairTest Executive Director, p. 2.

A Call to Action

Breaking Down the Barriers

Standardized multiple-choice tests have become America's cradle-to-grave arbiter of social mobility and a dominant force in shaping education across the nation. From nursery school entrance tests to multi-part licensing exams for the professions, students and workers are required to take tens of millions of standardized tests each year.

Yet a large body of evidence demonstrates that these exams fail to accurately and completely measure the aptitude and achievement of many individuals, particularly minorities and women. Too often tests are not properly constructed, validated, administered, or used. Because the tests are race, class, gender, language and culture biased, the result is the erection of unfair barriers which deny equal opportunity for education and employment to otherwise-qualified applicants.

The standardized testing explosion also has profoundly negative consequences for the quality of the nation's schools. Increasingly, curricula are narrowed to cover the

(continued on page 7)

How ACRE Stopped Testing

In 1987, the North Carolina legislature voted to stop requiring the administration of the California Achievement Test to first and second grade children (see Examiner, V. 1, #3). In the following article, Page McCullough, the Executive Director of the Atlantic Center for Research in Education (ACRE), explains how they organized this landmark victory. Since its founding in 1978 as a group of parents, teachers and teacher educators focusing on the needs of poor, minority and handicapped children in North Carolina's public schools, ACRE has actively worked against the misuse of standardized tests.

The North Carolina campaign to stop norm referenced standardized testing in the first and second grades began five years ago when the state's General Assembly voted to start using the California Achievement Test (CAT) in the early grades. ACRE opposed the use of the tests for many reasons: normed group achievement tests have low reliability and content validity for young children; the scores provided little useful information to teachers; the curriculum was becoming "test-driven"; and too much time was spent on testing.

Our work to repeal this mandate has

(continued on page 12)

CONTENTS

Weiss Letter 2
Scholarship Reforms 3
NY Bias Commission 3
Biased 'Merit' Awards 4
ETS Loses ERIC/TME 4
Teacher Tests 5
Holmes Opposes Tests 5
Harvard SAT 'Study' 6
Conference Schedule 8
NAEP Limits 11
K-12 Test Fraud 11
Testing Children 13
ETS Report 14
GATB Study 14

FairTest Conference June 17-18

TM017189

FairTest

National Center for Fair & Open Testing

Dear Friends,

Four years ago, Bob Schaeffer and I set out to revitalize America's movement for fair, open and pedagogically sound means of evaluating the nation's students and workers. Today we have succeeded far beyond our initial dreams.

My role as an "agent of change" has been fulfilling and demanding. Yet it is now an opportune time, for both me and FairTest, to change my role in the organization. Accordingly, around the time of our June National Conference, I shall step down as Executive Director. I will continue an active affiliation with FairTest by remaining as its President and serving as Senior Policy Consultant.

The FairTest Board has opened a national search to find a new Executive Director. My replacement will have the pleasure of working with FairTest's superb board, crackerjack staff, and national network of supporters. It is a great job, with unlimited potential. We would greatly appreciate any suggestions about potential candidates.

What's next for me? I am waiting for something interesting to emerge. If you know of any quixotic projects that need an organizer, please get in touch.

I hope to visit with you during our June conference.

John Weiss

John Weiss,
Executive Director

The National Center for Fair & Open Testing (FairTest)
P.O. Box 1272, Harvard Square Station
Cambridge, MA 02238; telephone (617) 864-4810.

Staff: *Mike Friedman*, Assistant; *Jim Gray*, Washington Liaison; *Noe Medina*, K-12 Consultant; *Monty Neill*, Managing Director; *Bob Schaeffer*, Communications; *Sarah Stockwell*, Sex Bias Coordinator; *John Weiss*, Executive Director.

We would like to thank the Ford Foundation and the Rockefeller Family Fund for supporting our campaigns to ensure that the tens of millions of standardized tests annually administered to students and job applicants are fair, open and educationally sound.

EXECUTIVE DIRECTOR

RESPONSIBILITIES: Play lead role developing strategies, implementing advocacy programs, educating the public and raising funds. Represent organization in national arena; supervise staff and consultants; plan action agenda; and oversee resource allocation. Report to Board of Directors.

QUALIFICATIONS: A track record of leadership effectiveness in a planning and management position is crucial. Experience with civil rights, feminist, education reform or citizen action issues is desirable. Public speaking, writing, fundraising and organizing skills are a plus, along with knowledge of testing issues. Must be willing to travel. FairTest is an affirmative action employer.

SALARY: $35,000 to $40,000 annually with health insurance and four weeks vacation.

STARTING DATE: June 15, 1988 or as soon thereafter as is mutually agreeable.

TO APPLY: Send resume/curriculum vitae, writing samples and names of three references to Board of Directors, FairTest, PO Box 1272, Cambridge, MA 02238.

CLOSING DATE: Applications must be received in the FairTest office by 5 p.m. on Monday, May 2, 1988.

Resignations

ACT President Olugh Davidsen and ETS Vice-President Robert Solomon have both recently announced their resignations. To date, neither has applied to be FairTest's Executive Director.

Correction

In the last *Examiner*, we encouraged readers to order the *Special Issue of the Negro Educational Review: Testing African American Students*. We made a mistake on the price. The book-length special issue costs $20 (Box 2895, General Mail Center, Jacksonville, FL 32203).

Scholarship Reforms Narrow Gender Gap in New York Board of Regents Awards

A new procedure that reduces the role of test scores in choosing winners has greatly increased the number of females receiving the 25,000 college scholarships annually awarded by New York's Board of Regents. Under the old formula, based solely on Scholastic Aptitude Test (SAT) or American College Test Assessment (ACT) results, boys had received 60% of the basic Regents Scholarships and 72% of the top "Empire Excellence Awards" worth more than $10,000 each. The 1988 reforms, in which classroom grades and test scores were weighted equally, produced a 50-50 gender split in Regents Scholarships and hiked girls' share of Empire Awards to 40%.

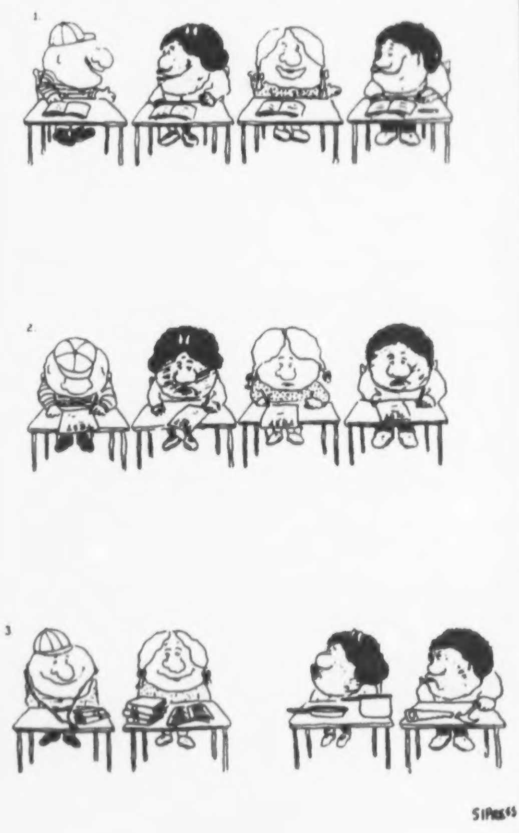
Though data on the ethnic background of winners has not yet been released, a preliminary analysis indicates that fewer scholarships were awarded this year to students from upper-income high schools. According to Blair Horner, Legislative Director of the New York Public Interest Research Group (NYPIRG) which coordinated the coalition of civil rights, feminist and consumer organizations that fought for scholarship reforms, "These results are a step in the right direction. Even the College Board admits that high school performance is a better predictor of ability to do college work than test scores. That's why grades should be the major factor in any fair scholarship competition. But women and minorities are still forced to play on an uneven field when tests scores are used in any manner."

New York Selects Test Bias Study Commission

Legislative leaders in Albany have appointed the ten members of the state's first-in-the-nation panel charged with examining statistical data on bias in college and university admissions tests. The panel will recommend ways to make the exams more fair (see *Examiner*, Vol. 1, Nos. 2 & 3). The New York commission includes a number of strong supporters of the testing reform movement as well as several members with ties to manufacturers of standardized, multiple-choice exams.

The panel, which begins meeting this Spring, includes:

- **Richard Barrett**, Consulting Psychologist, Adjunct Professor of Psychology, New York University
- **John Dovidio**, Assoc. Professor of Psychology, Colgate University
- **Jay Hershenson**, Vice Chancellor for University Relations, City University of New York
- **Winsor Lott**, former Director of Educational Testing, New York State
- **Lilliam Malavé**, Asst. Professor of Education, State University of New York at Buffalo
- **Jason Millman**, Professor of Educational Research Psychology, Cornell University
- **Martin Shapiro**, Professor of Psychology, Emory University
- **Maria Torres**, Asst. Professor and Director, Bilingual & Bicultural Ed. Program, Columbia Teachers College
- **Arnold Webb**, Asst. Commissioner of Education Programs, New Jersey
- **Napoleon Williams, Jr.**, Assistant Counsel, NAACP Legal Defense and Educational Fund, Inc.



Biased Tests Still Cost Women, Minorities Millions in 'Merit' Scholarships

Despite growing pressure from testing reformers and advocates of equal educational opportunity, the 1988 National Merit Scholarship competition has once again produced lopsided results in favor of boys. A FairTest analysis found that of this year's 15,370 semifinalists (the only students deemed eligible for Merit Scholarships), more than 60% are males and just 36% females (the gender of 4% could not be determined from their names). That's essentially no change from last year when 61% of the semifinalists and 62% of the scholarship winners were male. Though the National Merit Scholarship Corporation refuses to release ethnic breakdowns, it's likely that Blacks, Hispanics and Native Americans are also grossly underrepresented in the semifinalist pool.

This continued underfunding of scholarships for women and minorities stems from National Merit's refusal to change its policy of using Preliminary Scholastic Aptitude Test (PSAT) scores as the sole criterion for selecting semifinalists. Like the SAT, the PSAT contains biases against minorities, women and rural students which exacerbate already existing inequities, resulting in substantial score gaps between upper middle-class Whites and many minority groups and especially between White males and Black or Hispanic females. To compensate for some of the test's flaws, National Merit already sets geographic quotas and doubles the verbal score "to get a better mix of girls and boys" since the differential between males and females is less on the verbal section of the test.

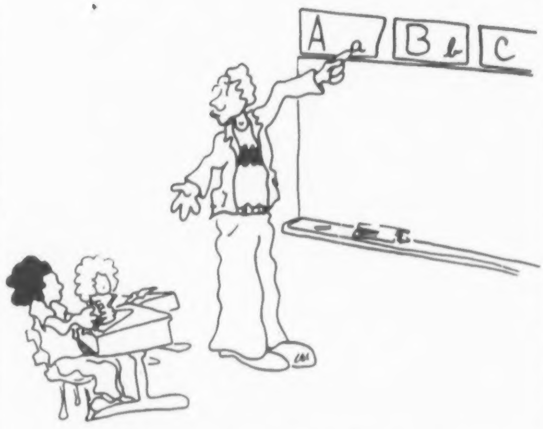
But such actions don't go far enough. Since test-makers agree that high school performance, not test scores, is the most accurate measure of ability to do college work, there's no excuse for continued reliance on an unfair screening device. If the goal of the Merit Scholarship competition is to reward the nation's most able students, it should immediately stop using test scores to select Semifinalists.

Fox Loses Control of Chicken Coop

The federal government has denied Educational Testing Service's application to continue operating the federal government's Educational Research Information Clearinghouse on Tests, Measurement and Evaluation (ERIC/TME). The action comes after a FairTest-led coalition of law professors, researchers and civil rights activists presented evidence that ETS used its control of the clearinghouse to unfairly promote its products. FairTest also claimed it was a violation of federal conflict-of-interest laws for the nation's largest testing company to operate the official federal information center on the products they market.

While ETS claimed that FairTest's charges had nothing to do with the government's decision not to renew the contract, an internal ETS memorandum blamed the "recent conflict-of-interest criticisms voiced by FairTest." ETS had operated the clearinghouse since its inception 17 years ago.

American Institutes for Research, of Washington, D.C., will run ERIC/TME under a $275,000 a year contract renewable until 1992. FairTest will serve as an affiliate for the new clearinghouse which will be run by former Dept. of Education researcher Dr. Lawrence Rudner.



"And after you learn the ABCs, we start on the CTBSs, the SATs, and the GREs"

ETS News

ETS-Princeton Review Suit Settled

A lawsuit against the Princeton Review (PR) SAT coaching program by Educational Testing Service (ETS) has ended with both sides claiming victory. In a consent agreement, PR agreed to pay ETS $52,000 in copyright infringement damages for use of ETS' questions in the PR course. PR President John Katzman claimed victory because ETS spent $1.2 million on the suit but settled for $52,000, and all other parts of the suit were dismissed. PR also gained the right to license from ETS as many questions as needed for their course; previously there had been a limit of 100 questions.

ETS Pulls NTE from NYC

Educational Testing Service (ETS) has stopped sending scores on its teacher test, the NTE, to New York City because the local school board incorrectly set cut scores on the test. ETS requires users of the test to validate its use. To restore the exam, NYC can either adopt the state's cut-off scores, which were set following a controversial ETS-run validity study of its own product (see *Examiner*, V. 2, #1) or do its own validation study.

ETS Halts NTE Use for Rhode Island Social Workers

In response to a complaint filed by the Rhode Island chapter of the American Civil Liberties Union, ETS has told the Rhode Island Department of Education to stop using the NTE to license school social workers. ETS said a new validity study must be conducted before the NTE may be used to certify anyone other than classroom teachers.

Where the Money Goes

What happens to ETS' booming revenues? Public reports from the firm given one answer. Executive salaries are skyrocketing. ETS President Gregory Anrig was paid $130,499 in 1984 (plus $28,410 in benefits). Top VP Robert Solomon was close behind with $115,896.

NES Teacher Tests Shot Down in Alabama and Georgia

Major lawsuits against teacher tests manufactured by National Evaluation Systems (NES) of Amherst, Massachusetts were recently settled on terms very favorable to minority educators in two southern states. NES is the nation's second largest distributor of teacher licensing exams.

In Georgia, the local National Educational Association affiliate won several significant reforms in return for ending its litigation charging that state-mandated certification and recertification tests were racially-biased and not a valid predictor of job performance. As part of the settlement, the state agreed to revise its National Evaluation Systems-produced tests and provide free study courses to teachers who have not yet passed them. In addition, the state will pay $6,000 in the form of a "study grant" to each teacher who lost certification as well as $2,000 to every provisional teacher who has not yet passed and enrolls in the two-week test preparation course. Further development of the Georgia teacher test will be done by Georgia State University, not by NES.

The six-year-old suit against Alabama's NES-designed teacher licensing test has also ended. In an initial settlement in early 1986, the state agreed to create new exams using a modified "Golden Rule" bias-reduction procedure, to stop using tests as the sole certification criterion, to bar testing of materials not in the official curriculum, and to allow plaintiffs to monitor the settlement. The state then tried to renege on the agreement, but a federal appeals court ruled in November, 1986, that the agreement was binding (*Examiner*, V. 1, No. 1). After months of unsuccessful legal maneuvering in U.S. District Court, the state finally accepted the original settlement and agreed to pay $500,000 to be divided among 900 or more Blacks who initially failed the test but passed after its redesign. The Alabama teacher test is now out for bids.

Northeast Holmes Opposes Teacher Tests

At its quarterly meeting on January 29, 1988, members of the Steering Committee of the Northeast Holmes Group, comprised of representatives from major research universities having teacher education programs, passed a resolution opposing the use of all current teacher tests. The resolution said that because the use of cut-scores from such exams to certify teacher candidates has a "highly disparate impact on Black teacher candidates," and there is little evidence of the validity of the exams "to differentiate between teachers who are competent and those who are not,"

"Be it resolved that the Steering Committee of the Northeast Holmes Group recommends against the use of such cut-scores on teacher tests until such time as evidence is available to show that use of such cut-scores differentiates among teacher candidates more on the basis of some independently measured indicators of teacher quality than on the basis of race."

EDUCATION WEEK

Volume VII, Number 25 · March 16, 1988

Harvard Study of Courses On S.A.T. Said 'Suspect'

To the Editor:

The Harvard University report that claims coaching for the Scholastic Aptitude Test does not help students raise their scores is suspect in several important respects ("Harvard Study Casts Doubt on Effect of Test Coaching," Research and Reports, *Feb. 24, 1988*).

First, the study was conducted by Dean K. Whitla, who is not only director of the university's office of instructional research and evaluation but also a trustee of the College Board, the firm that earns $70 million annually selling the S.A.T. and related products.

This is a blatant conflict of interest, akin to having Lee Iacocca do a study on what cars Harvard should buy for its campus police fleet.

The College Board and its hired guns naturally want to disparage coaching. If coaching works, then the well-documented racial and gender biases of the S.A.T. would be compounded by the inability of some students to pay for test-preparation courses.

It would also prove that the exam measures nothing more than test-taking skills. No wonder the College Board itself has written, "If the board's tests can be regularly beaten through coaching, then the board itself is discredited."

Second, the methodology of the study was poor. Incoming students were asked whether they had been coached. Such self-reporting is a notoriously unreliable survey technique. Many freshmen may have thought it unwise to tell Harvard they had been coached.

Third, the claim that those who admitted being coached did not show score increases much different from those of students who said they were not coached is contradicted by the facts. Harvard says those who were not coached showed gains averaging 67 points, while those who were gained 94 points.

This 27-point difference is highly significant as an average for a large group of people. The S.A.T.'s margin of error, within which the 27 points supposedly fall, applies only to individuals, a fact well known to any quantitative researcher in this field.

Ironically, when the national-average scores of minorities rise a point or two, the College Board touts this as an important gain.

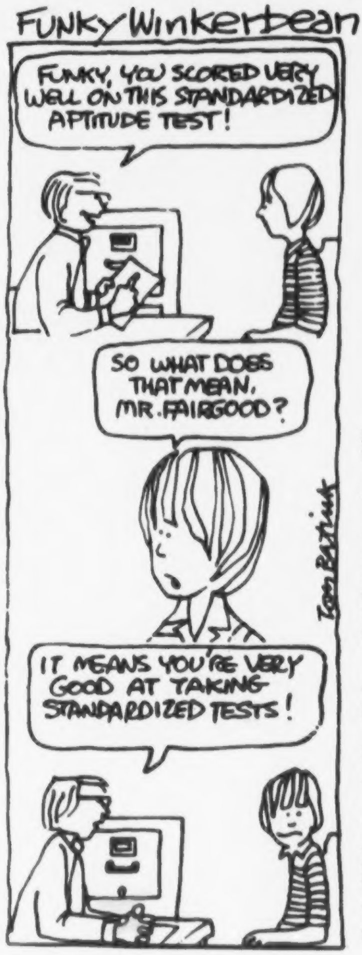
Fourth, the gains of both groups are inconsistent with other College Board claims. The board says the expected increase from junior to senior year for initially high-scoring students is only around 10 points. Something is wrong with either the Harvard data or the test-maker's claims.

Half a dozen recent independent studies all conclude that well-run test-preparation courses can increase scores by 100 points or more.

If the College Board really wants to test its claim that coaching does not work, it should join in co-sponsoring a blue-ribbon panel to independently study S.A.T.-preparation courses.

FairTest—the National Center for Fair and Open Testing—asked the board's president, Donald Stewart, to do just that a year and a half ago. We're still awaiting his response.

John G. Weiss
Executive Director
National Center
For Fair and Open Testing
Cambridge, Mass.


Funky Winkerbean
FUNKY, YOU SCORED VERY WELL ON THIS STANDARDIZED APTITUDE TEST!
SO WHAT DOES THAT MEAN, MR. FAIRGOOD?
IT MEANS YOU'RE VERY GOOD AT TAKING STANDARDIZED TESTS!
Tom Batiuk

Funky Winkerbean by Tom Batiuk · Field Enterprises, Inc 1979

More Proof that Coaching Works

The San Francisco Unified School District reported in September, 1986, that 76 predominantly low-income and minority students who received SAT coaching from Dr. Gary Gruber obtained average score increases of 44 points on the verbal and 67 points on the math portion of the test, a total gain of 111 points. The students were first tested on the November 1985 SAT, took the course, and then took the December 1985 SAT.

Federal Commission to Examine Test Validity

As one of its final acts under Reagan administration appointees, the U.S. Commission on Civil Rights will conduct a study on the validity of standardized tests. The Commission says it will look at the impact of tests on racial, ethnic and gender groups to determine whether their validity is scientifically up-to-date and in accord with civil rights laws. Regulations and case law on testing in education and employment settings will also be examined. The study was fueled by a desire of some Reagan appointees to find a way to overturn the California *Larry P.* decision (see *Examiner*, V. 2, #1). The project is expected to last 18 months.

Breaking Down the Barriers: A Call to Action

(from page 1)

limited subject areas multiple-choice tests can measure. In many districts, higher-order educational goals are subordinated to simplistic test score competitions. The study of literature and the practice of writing in English classes, for example, give way to memorizing analogies and taking multiple-choice vocabulary quizzes. This narrowing of the curriculum tends to affect low income and minority students most heavily.

Though test manufacturers and the agencies which mandate exams often restrict release of information on racial and gender performance, validity and educational impact, data obtained through lawsuits, Freedom of Information requests and limited Truth-in-Testing requirements present an alarming picture:

Kindergarten through High School

- Development and placement exams are becoming increasingly common for young children. In several southern states, tests determine kindergarten eligibility of five and six year olds. In 1984, 40% of the Black five-year-olds in South Carolina failed the state's "Ready-or-Not" exam and were tracked into remedial classrooms before they even began formal schooling.
- In 15 states students must pass standardized tests in order to be awarded high school diplomas, eight require tests for promotion, and 37 use standardized exams as part of a state assessment program. Nationally, Black, Hispanic and low-income students fail these exams at rates many times greater than middle and upper-income Whites.

College and University Admissions

- Hundreds of colleges use the Scholastic Aptitude Test (SAT) or the American College Test (ACT) to establish arbitrary cut-off scores below which applicants will not be considered. More than 750 government and private scholarship agencies also rely on test scores to help select winners. This presents a formidable barrier to Blacks and Hispanics whose SAT scores average 200 points below those of Whites and to women who score nearly 60 points lower than males despite earning higher grades in both high school and college.
- Standardized tests are routinely used by colleges to determine eligibility for entry into special programs. The University of Texas system, for example, requires that applicants for teacher training programs take the Pre-Professional Skills Test (PPST). Three-quarters of White applicants passed this entry hurdle, but only one-quarter of Black and one-third of Hispanic test-takers qualified in 1986.
- Additional multiple-choice tests are required for admission to graduate and professional schools. Again, minorities and women are at a profound disadvantage. For example, in a former version of the Law School Admissions Test (LSAT), Blacks averaged more than 100 points lower than Whites even when they obtained similar undergraduate grade point averages at identical institutions.

Occupational Licensing Tests

- Multiple-choice, standardized licensing exams are required by nearly 100 trades and occupations in the United States, and the number grows every year. Many professionals must take additional standardized tests to maintain their credentials. On most of these exams, the passing rate for minorities is significantly lower than for Whites.
- Seventeen states use the National Teachers Examination (NTE) to license classroom educators, and all but a few states require some exam for teacher certification. A study done in New York found that Whites were twice as likely to pass the "General Knowledge" section of the NTE as Blacks or Puerto Ricans. Females were twice as likely to fail as males.
- A host of lawsuits clearly illustrate how test bias can limit the opportunities of "blue collar" minority workers. For example, the question "What does R.S.V.P. signify?" was recently removed from the Washington D.C. police promotion exam after a lawsuit claimed such test questions unfairly discriminated against Blacks.

Reform Principles

FairTest was formed to address unfair barriers like these. The organization's agenda is shaped by four basic principles:

- *Tests must be fair and valid;*
- *Tests should be open;*
- *Tests should be viewed in their proper perspective; and*
- *Alternative assessment instruments must be developed.*

Test reformers have won many victories but much more remains to be done.

FairTest Conference

Since its founding in October, 1985, FairTest has led the national campaign to reduce the discriminatory impact of tests, open up the testing process, and educate decision-makers, the press and the general public about the problems of standardized, multiple-choice exams. Now it is time to broaden our movement by launching new initiatives to overhaul or eliminate tests which damage equal access and quality education.

FairTest's June 17 & 18 national conference "Breaking Down the Barriers" offers the entire testing reform movement an opportunity to share the latest victories, learn about the newest bias-reduction techniques and develop winning strategies. Look at the exciting agenda on the following pages. Then use the registration form on the back cover to sign up today.

See you in Washington in June.

National Testing Washington, D.C.

Breaking Down

Friday, June 17
Capitol Hill Reception

Caucus Room
Cannon House Office Building
Featuring Rep. Patricia Schroeder

Rep. Schroeder will address the general problem of test bias and misuse and the need for reforms to assure tests are fair, valid and properly used.

Wine and cheese will be served.

Saturday, June 18
Presentations and workshops

Howard Law School
2900 Van Ness St., N.W.
Washington, DC

Morning

9 a.m. — ***Coffee and Pastry***

9:30 a.m. — ***Welcome:*** Dean J. Clay Smith, Howard Law School

FairTest Executive Director John Weiss will speak on the history and goals of the testing reform movement.

Three short talks on reform victories:

Attorney Don Watkins on teacher tests in Alabama; Page McCullough on grade school tests in North Carolina; Blair Horner on truth-in-testing in New York.

11 a.m. — ***Simultaneous sessions on major themes.***

Speakers include:

Grade School Testing:
Harold Dent, Assn. of Black Psychologists & *Larry P.* case; David Feldman, Project Zero & Tufts University; Ann Cook, Community Studies, Inc., New York City.

University Admissions:
Denise Carty-Bennia, Northeastern Law School; Ron Vera, Esq., former MALDEF staff; David White, Testing for the Public, Berkeley, Cal.

Employment Exams:
Beverly Cole, NAACP Education Director; Bill Robinson, Lawyers Committee for Civil Rights Under Law; Diana Pullin, Dean, Boston College School of Education; Bob Schaeffer, FairTest.

12:15 P.M. — **LUNCH**

To register for the Conference, fill in and tear off the back page and send it

Reform Conference June 17-18, 1988

the Barriers

Afternoon

1:45-4:30 p.m. **Workshops**
Round 1 (1:45-3 p.m.)
Round 2 (3:15-4:30)

Grade School Testing:
Susan Bredekamp, National Association for the Education of Young Children; Ann Cook, Community Studies, Inc.; Monty Neill, FairTest.

Alternatives to Testing Schoolchildren:
David Feldman, Tufts U. & Project Zero; Harold Dent, Assn. of Black Psychologists.

Organizing A Test Reform Campaign:
Page McCullough, ACRE (see article, page 1).

Gender and Race Bias in National Scholarship Programs:
Denise Carty-Bennia, Esq., Northeastern Law School; Sarah Stockwell, FairTest.

Exploring the Gender Gap on the SAT:
Phyllis Rosser, Equality in Testing; Jim Loewen, U. of Vermont.

Language, Culture and Testing:
Orlando Taylor, Howard University; Guadalupe Valdez, U. Cal-Berkeley.

Using the Coaching Controversy to Force Test Reforms:
Beverly Cole, Education Director, NAACP; David White, Testing for the Public.

Problems of Teacher Testing:
Walt Haney, Boston College; Jim Gray, Lawyers Committee for Civil rights Under Law; Elaine Witty, Dean, Norfolk State U., School of Education.

Litigating Teacher Tests:
Al Kaufman, Esq. MALDEF, Texas; Bill Robinson, Esq., Lawyers Committee for Civil Rights Under Law; Don Watkins, Esq., Alabama.

NAEP and National Tests:
Marcella Dianda, National Education Association; Martha Jimenez, MALDEF; Noe Medina, Education Policy Research.

Auditing High Stakes Testing:
George Madaus, Boston College, former President, National Council for Measurement in Education; Ron Vera, Esq, former MALDEF staff; Vito Perrone, Carnegie Foundation for the Advancement of Teaching.

Employment Tests: Can They Be Fair and Relevant?
Richard Barrett, Consulting Psychologist.

Additional speakers and workshops will be announced.

4:45 p.m. **Closing**
Vito Perrone, Carnegie Foundation for the Advancement of Teaching

5-7 p.m. **Refreshments**

---

to FairTest with your registration fee. Further information will be sent to you.

The Evening Sun
Baltimore
March 21, 1988

The folly of comparing education test results

Our scores are better than your scores

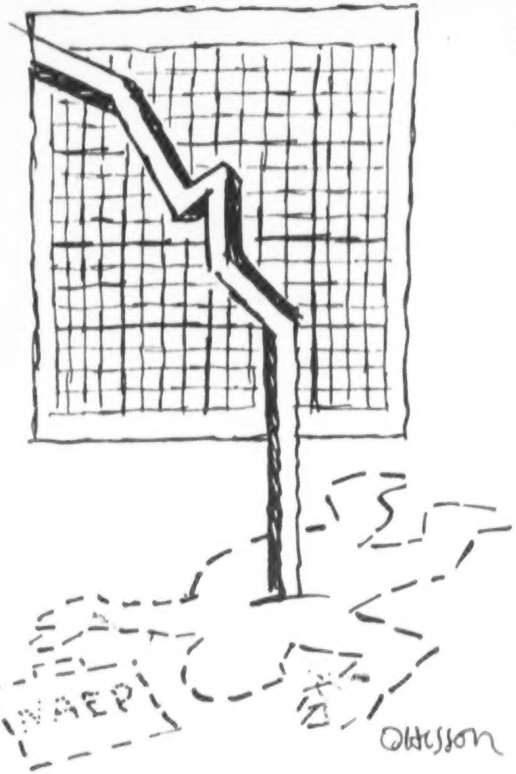
John G. Weiss & Arnold Fege

INFORMATION CHANGES behavior. What the government measures and publicizes becomes important to most Americans. When auto dealers were required to publish miles-per-gallon ratings, consumers started paying attention to fuel efficiency. When cigarette companies labeled tar and nicotine levels, smokers altered their behavior: today over half the cigarettes consumed in the U.S. are low in tar and nicotine, compared with 4 percent before labeling.

But publishing misleading information can lead to harmful changes in behavior. That is what many educators fear in the recent passage by the U.S. Senate, without hearings, of a bill to expand the National Assessment of Educational Progress (NAEP), the federal government's major public school testing program. By removing a ban on state-by-state and district-by-district comparisons of NAEP scores, the legislation will result in federal statistics purporting to compare the performance of American schools.

The consequences of the Senate proposal are predictable. Newspapers will rank-order states by their NAEP scores. Elected public officials — and candidates running against them — will promise to boost NAEP scores. Educators pressured to increase scores will teach to the test. Textbooks will be redesigned to focus on the limited topics NAEP covers. Class time will be spent preparing students for the tests. "Creative" measures might be taken to encourage certain low-scoring students to be absent on test days.

Under the proposed Senate plan, up to 700,000 students could be tested annually in civics, geography, history, math, reading, science and writing. Given the size and scope of this testing program, the NAEP exams will rely almost exclusively on the multiple-choice format. While such questions can be scored by a machine, they only measure the most rudimentary skills. The history test will ask about the dates of events, not their importance or meaning. The writing exam will test grammar and vocabulary, not communication or composition skills.



Today, many states and cities publicly rank schools by their reading test scores. The result is that reading instruction in these jurisdictions has come to mirror the tests. Class time is spent answering multiple-choice questions about short reading passages taken out of context. Students — especially our very brightest — become bored and alienated.

Instituting a national competition to boost multiple-choice test scores will magnify these problems a hundred-fold. NAEP exams will end up directing what is taught in our classrooms. George Madaus, past president of the National Council of Measurement in Education, recently noted that: "Whoever controls those powerful tests will control to a large extent what is taught and learned in American schools."

The plan to expand NAEP comes at a time when many other institutions and policy making bodies are starting to critically examine the impact standardized tests have on curriculum and society. Harvard Business School has dropped all admissions testing requirements. Bates, Bowdoin and Union College no longer require applicants to submit SAT scores. North Carolina legislators have voted to stop testing kindergarten, first and second grades. And a federal court has barred California from using IQ tests to place black children in programs for the "Educably Mentally Retarded."

The bill to expand NAEP has not yet been considered by the U.S. House of Representatives. Before endorsing the Senate measure, the House should hold hearings to obtain public comment about the implications of instituting a national exam.

What gets tested and compared communicates to educators and the community at large what is valued. At a time when commission after commission urges American schools to focus on higher order critical thinking skills, we risk reducing education to what a multiple-choice exam can measure. Without safeguards, the NAEP exams will become American education's "tail that wags the dog."

John G. Weiss is executive director of the National Center for Fair and Open Testing (FairTest), a Cambridge, Mass., organization concerned with the growing impact of standardized exams in schools and employment. Arnold F. Fege is director of government relations for the National PTA.

House and Senate Limit NAEP Expansion

House and Senate conferees have apparently reached agreement on the proposed expansion of the National Assessment of Educational Progress (NAEP) included in H.R.5, the omnibus education bill. The committee compromise includes a number of victories for the test reform coalition including the PTA, FairTest and other civil rights and education groups concerned that NAEP not become a national test that could drive US public education.

While NAEP will be expanded via two pilot projects and state-by-state comparisons will be allowed, such comparisons will not be mandatory, and states can opt in or out of them. Reports will be made to Congress on each pilot project, allowing public oversight before any further developments. In addition, the NAEP policymaking board will be composed entirely of new members, making it independent of ETS, the producer of the NAEP tests.

"If more changes are proposed for NAEP, Congress should hold full public hearings."--Arnie Fege, PTA

NAEP expansion had been proposed in the Senate by Senators Kennedy (D-Mass) and Hatch (R-Utah) and Secretary of Education William Bennett. Before the bill even reached the conference committee, the reform coalition initiated by FairTest had already won a number of safeguards including use of bias reduction techniques, "truth-in-testing," an independent NAEP staff and a racially diverse policymaking board (see *Examiner*, Vol. 2, No. 1).

But the coalition remained concerned that state-by-state comparisons would have a long-term negative impact on education and continued to discuss the issue with Senators, Representatives and staff members. The conference committee ultimately agreed that the potential dangers were severe and endorsed only a limited expansion.

Under the final NAEP plan, larger assessments of third-grade reading will be conducted in 1990 and of math in grades 3, 7 and 11 in 1992. States can decide whether or not to be part of the expanded testing and once the testing is completed, whether or not to have their state's scores reported separately. The legislation will not be finalized until late April.

Arnie Fege of the National PTA concluded, "There is an influential minority that wants to use tests to help centralize educational decision making. A narrowly designed federal test, however, is not in the best interest of kids, of school reform or of assessing education. Once our coalition raised these issues, Congress understood them and made the correct decisions. But we need to be vigilant because this kind of legislation will be back again. If more changes are proposed for NAEP, Congress should hold full public hearings."

Problem in NAEP Reading Test

Release of the 1986 National Assessment of Educational Progress (NAEP) reading test scores was delayed more than five months because no one believed the apparent results: a steep drop in reading performance among 9- and 17-year-olds since 1984. No other assessment concurred with NAEP's finding. The test-maker, ETS, finally concluded the problem was not in the test but in comparing it with previous results. However, the experts still don't know why this "anomaly" occurred. As a result, the report, "Who Reads Best," contains no year-to-year comparisons.

The report concluded that many students do not read well. The authors observed, "Many school programs add to the problem by stressing basic skills for poor readers rather than comprehension and critical thinking." But they fail to note that the proliferation of standardized exams leads to the misplaced emphasis on basic skills, the negative consequences of which fall most heavily on low-income and minority children.

K-12 Test Fraud Charges Filed in All 50 States

Charges of consumer fraud have been filed in all 50 states against four major K-12 test manufacturers by Dr. John Jacob Cannell of West Virginia's Friends for Education. Cannell charged that the tests are a fraud because their results show that all states, 90% of the nation's school districts, and 70% of all elementary students score above average.

The allegedly fraudulent tests are the California Achievement Test, the Comprehensive Test of Basic Skills, the Stanford and Metropolitan Achievement Tests, the Iowa Tests of Basic Skills, and Science Research Associates' SRA exam. Cannell is demanding that the test publishers identify school districts that use their tests, update the norms without charge, and make compensation for misrepresentations.

How ACRE Stopped Testing in North Carolina

(continued from page 1)

been as much a state of mind as any set of techniques. We knew we were in for a long battle because many politicians and parents want accountability and they see these tests as a good way to get it. The new testing program also enjoyed the fervent support of then Governor James Hunt and State Superintendent of Public Instruction Craig Phillips.

In such an unfriendly atmosphere, ACRE began the campaign by researching the issue, monitoring relevant boards, educating parents and teachers, introducing ourselves and our position to the legislature, building a constituency, training citizens to lobby and "waiting for daylight."

ACRE's campaign efforts included monitoring meetings of the State Board of Education and the North Carolina Testing Commission, which is responsible for advising the State Board on all testing programs administered by the state. Our observers made it possible for us to know when questions were being raised about the CAT in early grades. When an elementary school principal asked for a review of the program, ACRE and the North Carolina Association for the Education of Young Children (NCAEYC) got on the agenda to present our views. We learned to hone our arguments, and we learned that the testing commission was never going to recommend a change in the program.

When the legislature was in session, we monitored relevant committee meetings and made our position known. Not more than a dozen legislators supported us at first, but we learned how the General Assembly worked and we became familiar faces.

ACRE educated parents and teachers about the issue whenever we could. We held workshops and published articles about the issue in our newsletter. We created our own "Parent and Citizen Test Review Commission," which publicized our position opposing the testing of first and second graders. With these efforts, we built a small group of dedicated teachers who were willing to learn to lobby.

ACRE found a firm ally in the North Carolina Association for the Education of Young Children. This organization has what our small organization does not: 1,700 dues-paying members. The NCAEYC is composed of professors, teachers and day-care providers concerned with the quality of services for children from infancy to 8 years-old. ACRE had experience lobbying and cutting through the bureaucratic maze of the State Department of Public Instruction; NCAEYC had voters in every district. There's nothing like a lecture from one's first grade teacher to change a legislator's mind!

In 1987, a member of the NCAEYC persuaded a respected and popular legislator, who had young children, to introduce a bill to stop testing altogether in the first and second grades. By this time we had a different governor, who was not interested in the issue, and legislators were beginning to hear complaints about the time testing was taking and the stress it caused young children.

ACRE and the NCAEYC's years of organizing paid off as teachers testified before legislative committees and our phone trees went into high gear. ACRE presented our position paper, signed by many elementary teachers. Our most persuasive arguments were that young children are poor test takers, so the results are not reliable, and that the test ignored very important goals of our curriculum which cannot be tested with paper and pencil tests. The unanimous vote for the bill in the Senate represented a major change in perspective by lawmakers.

However, the bill stalled in the House when opposition from Superintendent Phillips and former Governor Hunt surfaced. Our legislative allies then attached a provision to the education appropriations bill eliminating the requirement for testing, and this provision was enacted into law.

Since our initial success with the issue, the NC Association of Psychologists and the NC School Psychologists have joined with ACRE, the NCAEYC, and the North Carolina Association of Educators (NCAE) to form a larger coalition. ACRE continues to monitor and keep our network and key legislators informed. The proponents of testing are attempting to restore state-mandated use of the CAT in grades one and two.

Our campaign suffered from the usual lack of time and money. Six years is a long time to sustain a volunteer group and we are no match for the personnel at the disposal of the bureaucracy and the test companies. Our work also suffers from a lack of parental support, which is a serious flaw. Parents have every right to know about their child's progress in school and we have not done a good job of offering alternatives to the lousy measures now being used.

Our strengths in this campaign included an astonishingly persistent group of volunteers who became increasingly skilled in lobbying in a timely manner. We did our homework and knew a variety of arguments to use and with whom to use them. We were pragmatic and non-partisan and in the end had voters from the Black Caucus, liberal Democrats, and conservative Republicans. Finally, we owe a great deal to three experienced women lawmakers who were willing to work hard and trade chips for this issue.

Our conclusion is that a successful campaign must be prepared for a protracted battle, organize a coalition that can develop significant public support from critical sectors, diligently monitor relevant state bodies, find the persuasive arguments and evidence to influence decision-makers, and locate legislators willing to take the issue as their own. Finally, victories must be defended and hopefully expanded.

Page McCullough will conduct a workshop on ACRE's campaign at the FairTest conference in June.

The New York Times

Copyright © 1988 The New York Times NEW YORK, TUESDAY, MARCH 1, 1988

Why Testing Tots Is a Wrong Answer

To the Editor:

Before deciding to require all kindergarten students to pass a standardized examination in order to enter first grade (Education page, Jan. 13), New York City officials should first look at the North Carolina experience with similar tests. There, the State Legislature recently voted to stop testing kindergarten, first- and second-grade students.

North Carolina curtailed its testing program after learning that instruction was being "test-driven." Against their better judgment, teachers were altering what they taught to conform to the test. Yet, legislators also found that there was very little overlap between the content of their tests and the state's curriculum objectives. The days devoted to testing (and taking practice tests) were also taking up valuable instructional time.

Finally, state officials found their tests had very low reliability for young students. Student scores often fluctuated significantly from one test to another — yet the scores from a single test were being used to deny students promotional opportunities.

Recent research demonstrates that students who are held back — even in the first grade — learn less and are more likely to drop out of school than similar students who were promoted. Other studies indicate that orally administered standardized tests — such as those advocated by Richard R. Green, who becomes the New York City Schools Chancellor on March 1 — are often poorly administered and systematically underpredict the abilities of black and Hispanic youths.

Five-year-olds do not need to be drilled on how to give standardized answers to standardized questions. Instead, what New York City education officials should be focusing on are the strategies that will help kindergarten children to think and enjoy learning.

JOHN GORDON WEISS
Exec. Director, FairTest (National Center for Fair & Open Testing)
Cambridge, Mass., Feb. 9, 1988

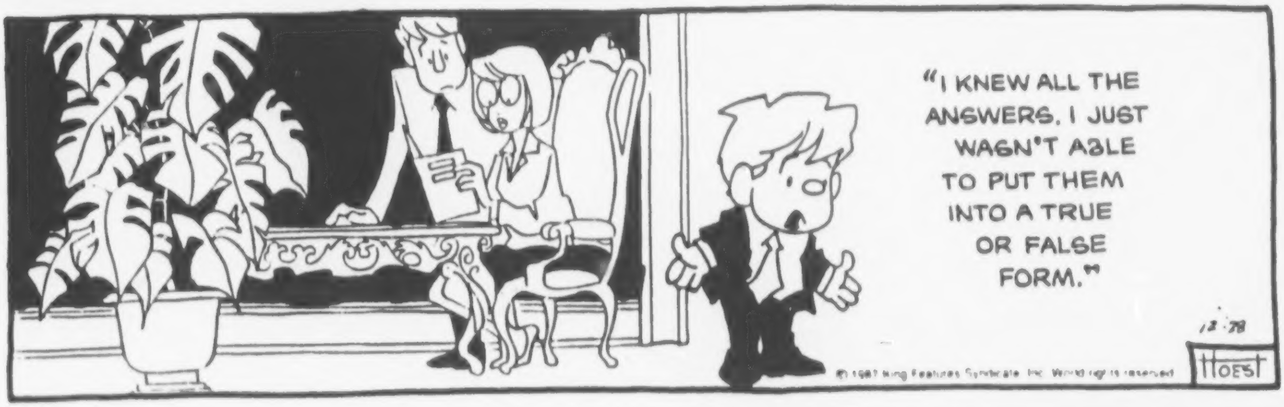
Testing Young Children

Though North Carolina just stopped requiring the test for first and second grades, the Georgia Board of Education has recently mandated the use of the California Achievement Test (CAT) for all that state's kindergarten children. CAT results will help determine whether a child is promoted to the first grade. Students who are not promoted will be placed in a remedial program.

Yet according to the National Association for the Education of Young Children (NAEYC), research demonstrates kindergarten retention is "educationally and psychologically harmful" and has "long-term negative effects." Susan Bredekamp of the NAEYC explained, "Our children are being tested too early and failed too soon."

In November 1987, the National Association of Early-Childhood Specialists in State Departments of Education condemned the use of transitional programs, delayed school entry, and repetition of kindergarten. The tests used for such placements, the group concluded, generally do not meet "acceptable standards of reliability and validity." The harmful effects they noted include racial, ethnic and soci-economic segregation, creating a sense of failure among very young children, and widening the age and ability dispersion within classrooms.

WHAT A GUY! / Bill Hoest


"I KNEW ALL THE ANSWERS. I JUST WASN'T ABLE TO PUT THEM INTO A TRUE OR FALSE FORM."
12·78
HOEST
© 1987 King Features Syndicate, Inc. World rights reserved

ETS: A Pretense of Accountability

Responding to growing concerns raised by feminist, civil rights, consumer and an increasing number of academic groups, the Educational Testing Service (ETS) devotes its "Trustees Public Accountability Report" in the firm's *1987 Annual Report* to charges of test bias. But rather than address the fundamental issues, ETS continues to evade them.

For the most part, the report consists of self-congratulatory descriptions of limited policies and procedures ETS was forced to institute by legally mandated reforms such as Truth-in-Testing and New York's Item Analysis Study Commission.

Still, ETS does admit that "the impact and ramifications of cultural and linguistic differences on test performance are not yet fully known," and that the SAT does underpredict how well women will do in college. But then come the excuses: The "question of test bias is extraordinarily complex"; "no two individuals are identical, and so it is with groups," etc.

Acknowledging that "it is possible...for tests to be biased," ETS takes great pride in two procedures which it says "are central to ensuring that its tests are as free of bias as possible." One is "sensitivity review" which consists of panels of readers screening questions for surface defects. But it is just such committees which have failed to remove blatantly biased questions in the past.

The second is what ETS calls "differential item functioning", a statistical procedure in which members of different groups, such as Blacks and Whites, are matched according to their "relevant knowledge and skill." Questions on which the two groups differ are then "flagged."

There are two problems with this procedure. First, the measure ETS uses to match groups by "ability" is their overall score *on the entire test*. Of course, this assumes the test, as a whole, is bias free. Second, even after the procedure identifies an item with a high "differential", ETS doesn't necessarily discard it. Instead, it goes back for further review by the same type of committee which approved it in the first place. Unlike the reforms based on the Golden Rule procedure, there's no guarantee that "high differential" questions are ever removed.

Even more disingenuous is ETS' "overriding concern" expressed at the end of the accountability report. ETS trustees warn that if "unspecified" test bias is blamed for score differentials, people will overlook the educational, social and economic factors that contribute to score gaps. Then, "the affected individuals and groups--and society as a whole--would pay dearly for such an oversight."

The fact is that current standardized tests tend to *exacerbate* social inequities. Thus, tests like the SAT make the problem even worse for minorities, women and students from low-income families who already pay an unacceptably high price.

ETS' spirited defense is hardly surprising: Testing is a booming, extraordinarily profitable business. Since 1983, ETS' income from testing activities has increased more than 50%. In the last two years alone, ETS has logged more than $18.5 million in "excess revenue."

The Tail Wags the Dog

California's incentive program, "Cash for CAPS," in which schools receive funding based on their students' average scores on the California Assessment Program, has produced some curious fallout. In some districts, students organized and threatened to boycott the exams unless the schools used the money for projects of the students' own choosing, on the theory they should reap the fruits of their test-taking labors. Elsewhere, some schools suspended the regular curriculum to teach to the test. Since the funds are dispersed by a formula that weights the scores with socio-economic data, at least one school trained its teachers to coach students to respond to questions about family background so as to "lower the school's socio-economic status and thus raise their relative standing on the test norms." (*Policy Briefs*, No. 4, Far West Laboratory)

Feds Correct Jobs Test Race Bias

In recent years, many states have begun experimenting with a controversial system developed by the U.S. Employment Service to select prospective employees for job referrals based on their scores from the General Aptitude Test Battery. The GATB consists of 12 separately-timed exams designed "to predict an applicant's expected performance in *all* jobs" (emphasis added). Last year the GATB was administered to more than 650,000 job applicants across the nation, an annual figure which is expected to rapidly increase.

Because Blacks and Hispanics, on average, do not score as high on the GATB as Whites, the U.S. Employment Service has converted raw scores into percentiles within each ethnic group. Applicants are then ranked based on these within-group percentiles. Recently William Bradford Reynolds of the Reagan Administration's Justice Department claimed this procedure was racially discriminatory.

To investigate the "reverse bias" charge, the National Academy of Sciences (NAS) set up a special committee which recently released its "Interim Report: Within Group Scoring of the General Aptitude Test Battery". While the NAS committee "neither approves nor disapproves" the GATB or race-conscious employment practices, it did conclude, "that a top-down within-group referral rule is an effective way to balance the conflicting goals of productivity and racial equity. *This referral rule substantially reduces adverse [racial] impact with very little loss in the average predicted job performance of those referred*" (emphasis added).

The NAS committee continues to study the overall validity of the GATB as well as alternative job referral methods. (Committee on GATB, 2101 Constitution Ave., N.W. Washington, DC 20418)

Subscribe to the *FairTest Examiner*

The *FairTest Examiner* will keep you up-to-date on the growing movement for fair & open testing. You will be able to do more than just read about FairTest campaigns--you can help make them successful!

_____ *FairTest Examiner*, published quarterly...Individuals $15/year; Institutional checks $25

_____ *Standardized Testing Reform Sourcebook, 1988*, (200 pages)...$25

_____ *Sex Bias in College Admissions Tests: Why Women Lose Out*,
by Phyllis Rosser and FairTest staff, 2nd edition (40 pages)...$7.95

_____ *The Reign of ETS*, by Allan Nairn and Ralph Nader (550 pages)...$30

_____ *Beyond Standardized Tests: Admissions Alternatives that Work*,
by Amy Allina with FairTest staff (15 pages)...$5.50

Updates, issued bi-annually (spring & fall). $7/year/topic:

_____ *Testing Reform Initiatives*
_____ *K-12th Grade Testing*
_____ *University Admissions Tests*
_____ *Teacher Testing*
_____ *Employment Testing*
_____ *Sex Bias in Testing*
_____ *Test Coaching*

Order total _____

Tax-deductible contribution _____

TOTAL ENCLOSED _______
(enclose check to FairTest)

_____ **Back issues of *FairTest Examiner*** $4 each: Vol___, No.___ (all 4 for $15)

_____ **Moving.** I have moved/will be moving. My new address is (Please enclose old label or write old address.):

Name ______________________________ Title ______________________________

Organization ______________________________

Address ______________________________ Zip ______________

Make checks payable to FairTest and mail to: FairTest, P.O. Box 1272, Harvard Square Station, Cambridge, MA 02238

Breaking Down the Barriers

FairTest National Conference on Testing Reform

Friday, June 17 & Saturday, June 18, 1988 *Washington, D.C.*

Featuring Rep. Patricia Schroeder, Friday evening, Capitol Hill Reception

Presentations and Workshops, Saturday, Howard University Law School

See Pages 8 & 9 for Schedule of Events

Conference Registration

______ Yes, I will attend Saturday, June 18 at Howard Law School

______ I enclose $30 conference fee for Friday and Saturday events and lunch on Saturday

______ I prefer vegetarian food for lunch.

______ Send information about group hotel rates.

______ I am most interested in the following three workshop topics, in order of preference (this is not a workshop registration, but a "straw poll"):

1. ____________________
2. ____________________
3. ____________________

Daytime Telephone (   ) ______________

Please tear off *entire* back page with your address label and mail to FairTest with your registration fee.

FairTest
National Center for Fair & Open Testing
P.O. Box 1272
Harvard Square Station
Cambridge, MA 02238
(617) 864-4810

Bulk Rate
U.S. Postage
PAID
Cambridge, MA
Permit No. 53278

Journal Printing Company
Revere, MA 02151

Complimentary Copy.
Please Subscribe today.
See page 15.

Address Correction Requested

FairTest
EXAMINER
The National Center for Fair & Open Testing Volume 2, Number 3 Summer 1988

Breaking Down the Barriers

Sen. Albert Gore Jr. (D-TN) with FairTest Board Members (l. to r.) Chuck Stone, Bill Robinson and Denise Carty-Bennia.

More than two hundred civil rights leaders, education reform organizers, academic experts and public policy-makers, as well as a handful of testing company staff, gathered in Washington, D.C. on June 17 and 18 for FairTest's national conference "Breaking Down the Barriers". Participants came from more than half the states and a dozen national organizations reflecting the spreading breadth of the testing reform movement.

The weekend began with a keynote speech by U.S. Senator Albert Gore at a well attended Capitol Hill reception. The Tennessee lawmaker denounced the bias that pervades many exams noting that "standardized tests may serve to make existing [social] inequities worse" and concluding "Our progress won't be complete if education and professions are closed to talented individuals because biased exams prevent them from competing on a level playing field."

History and Goals

Saturday was devoted to sessions and workshops at Howard University Law School where experts and activists debated content, tactics, and priorities for the testing reform movement.

In his opening address, FairTest Board President and retiring Executive Director, John Weiss, reviewed the history and progress of campaigns to ensure that tests are fair, open and educationally sound.

Though proud of recent public policy reforms and public education activities that have demonstrated "the emperor has no clothes," he cautioned against self-satisfaction. According to Weiss, the testing reform movement faces an increasingly sophisticated industry that has diversified dramatically in the past decade. To counter the threat to educational quality and equal opportunity, Weiss said, FairTest is expanding its emphasis from university admissions tests to focus on the testing explosion in the grade school and employment arenas.

Victories

Following Weiss were presentations describing three recent campaigns which won significant testing reform victories: the Alabama law suit forcing an overhaul of the state's teacher test; ACRE's suc-

(continued on page 8)

CONTENTS

National Policy Panel 2
National Merit Challenged 3
University Admissions 4
NY Test Bias Study 4
ACT Under Fire 5
Grade School Testing 6
Arizona Reduces Testing 7
Segregation by Testing 7
Gore Conference Speech 8
Vito Perrone Remarks 9
Employment Testing 11
Feds Abandon PACE 13
Testing Explosion 16

FairTest Board Creates National Policy Panel

At the Washington conference, FairTest's Board of Directors announced the creation of a National Policy Panel, comprised of education, civil rights and public policy leaders, to advise the Board and staff on strategy and program development. The distinguished members of the new body include:

Leonard Beckum, Dean, School of Education, City College, CUNY, New York City.

Beverly Cole, Education Director, NAACP, Baltimore.

Harold Dent, Association of Black Psychologists, Berkeley, CA.

Arnie Fege, Director of Governmental Relations, National PTA, Washington, DC.

Howard Gardner, Project Zero, Harvard University, Cambridge, MA.

Asa Hilliard, School of Education, Georgia State University, Atlanta, GA.

Congressman Matthew Martinez, D-California

Deborah Meier, Principal, Central Park East Secondary School, New York City.

Sara Meléndez, Associate Director, Office of Minority Concerns, American Council on Education, Washington D.C.

Aurelio Montemayor, Intercultural Development Research Assoc., Chair, National Coalition of Advocates for Students, San Antonio, TX.

Harry Pachon, Executive Director, National Association of Latino Elected and Appointed Officials, Los Angeles, CA.

Alfredo de los Santos, Vice-Chancellor, Maricopa Community Colleges, Phoenix, AZ.

Eleanor Smeal, former President, National Organization for Women, Fund for a Feminist Majority, Arlington, VA.

Ted Shaw, NAACP Legal Defense Fund, Los Angeles, CA.

Gail Thomas, Dept. of Sociology, Texas A&M University, College Station, TX.

Congressman Theodore Weiss, D-New York.

A working session at the FairTest Conference.

The National Center for Fair & Open Testing (FairTest)
P.O. Box 1272, Harvard Square Station
Cambridge, MA 02238; telephone (617) 864-4810.

Staff: Mike Friedman, Administrative Asst./Publications; Mary Gilleece, Office Manager; Noe Medina, K-12 Consultant; Monty Neill, Managing Director; Sarah Stockwell, Gender/Race Bias Coordinator; Bob Schaeffer, Communications.

Board of Directors: Prof. Denise Carty-Bennia, Northeastern University Law School, Chair; Prof. Pamela George, NC Central University; Paul Pottinger, New Manhattan Corp.; William Robinson, Lawyer's Committee for Civil Rights Under Law; Bob Schaeffer, Treasurer; Steve Solomon, Esq.; Chuck Stone, Philadelphia Daily News; Ron Vera, Esq., Barbosa & Vera; John Weiss, President.

Partial funding for FairTest's work comes from the Ford Foundation and the Rockefeller Family Fund.

Conference photographs by Mike Friedman.

The Mark of the Beast?

What is ETS's address? Most knowledgeable observers know that the Princeton name on the firm's letterhead is nothing more than a public relations ploy. In fact, the test-maker's primary location has been on Rosedale Road in Lawrenceville, New Jersey, since 1958.

But why doesn't ETS use its proper street number, even on internal correspondence? A recent letter printed in an internal ETS publication and forwarded to FairTest sheds some light on the issue.

According to long-time employee Paul Holland, ETS's official address is *666* Rosedale Road. As Holland notes, "No benign numeral is 666, no artful arithmetic alliteration; it has been known since ancient times as the number of the beast, a sign of the devil! 'six hundred, three-score and six' (*Revelations* 13:18)."

Is it just coincidence?

Civil Rights, Feminist Leaders Challenge National Merit Formula

Testing for the Public's David White and NAACP Education Director Beverly Cole at FairTest Conference

A dozen and a half of the nation's major civil rights and feminist organizations and leaders have joined FairTest in calling on the National Merit Scholarship Corporation (NMSC) to stop using the Preliminary Scholastic Aptitude Test/ National Merit Scholarship Qualifying Test (PSAT/NMSQT) as the sole qualifying factor for semifinalist status for the scholarships it administers. The groups cited data gathered by FairTest demonstrating that college admissions tests such as the SAT, from which the PSAT/ NMSQT is derived, are biased on the basis of gender, race and class and fail to predict academic ability. (See *Examiners* Vol. 1, Nos. 1 & 2 and Vol. 2, Nos. 1 & 2.)

The signers, who include the National Association for the Advancement of Colored People, the Mexican American Legal Defense and Education Fund, the Asian American Legal Defense and Education Fund and the National Organization for Women, called for a new formula that would take into account high school performance. They also requested that the NMSC release the number and dollar amount of awards given to each racial/ ethnic group by gender. Finally, schools and corporations which fund scholarships through the NMSC were urged to develop alternative awards procedures if the NMSC refuses to change its formula.

The letter stemmed from the growing frustration with NMSC's lack of substantive response to evidence that merit scholarship winners are being determined unfairly. The PSAT/NMSQT is currently the sole factor used to determine semifinalists for 9000 scholarships worth $34 million that are annually awarded by the NMSC and the College Board through the National Merit Scholars Program, special corporate scholarships, the (Black) Achievement Program, and the National Hispanic Scholars Program.

For the past two years, FairTest has documented that a disproportionate number of scholarships are given to males and to Whites due to this misuse of test scores. In April, 1987, FairTest's report, *Sex-bias in College Admissions Tests: Why Women Lose Out*, showed that males win nearly two-thirds of National Merit Scholarships because young women comprise only about a third of the test-based semifinalist pool. Yet girls get better grades than boys in both high school and college.

This April, FairTest estimated that minority students, who constitute over one quarter of high school seniors, most likely win less than 15% of total scholarship dollars as a result of reliance on PSAT scores (See table).

The NMSC's responses to FairTest's claims have ranged from a misleading "our competition is an open competition... We have no quotas related to sex or ethnic origins," to a lame claim that test scores reflect a number of complicated factors, such as "boys take more advanced classes in high school, particularly in the math and sciences...."

But NMSC does have geographic quotas and doubles the Verbal scores in an attempt to get a better mix of boys and girls. Moreover, the College Board's own data show that even when boys and girls, whites and non-whites, take four, or even five years of math in high school, there is

(continued on page 4)

SAT AVERAGE COMBINED SCORES BY GENDER AND RACE
(400 to 1600 scale)

White M=969
Asian M=946
White F=912
Asian F=897
Native M=855
Mexican M=845
Puerto M=820
Native F=790
Mexican F=775
Black M=748
Puerto F=744
Black F=705

Asian=Asian/Pacific Americans
Native=American Indians
Puerto=Puerto Ricans
Mexican=Mexican-Americans

(from Ramist and Arbeiter, *Profiles, College Bound Seniors, 1985.*)

University Admissions Testing

NMSC *(from page 3)*

still a substantial SAT Math score gap.

Most telling was NMSC's eagerness last year to pass the buck when faced with evidence of bias against females in the PSAT. Marianne Roderick, senior vice president, told a reporter that the question should be asked of the test-maker (the Educational Testing Service). Apparently, the NMSC accepts no responsibility for ensuring the device they use for giving out millions of dollars in scholarships is free of bias against women and minorities.

Future action in the National Merit reform campaign will depend on the NMSC. To date, FairTest and its allies have received no response.

Signers of the public letter to the National Merit Scholarship Corporation board of directors include:

FairTest, The National Center for Fair & Open Testing
Asian American Legal Defense and Education Fund (AALDEF)
Center for Women Policy Studies
Equality in Testing Project, Phyllis Rosser, Director
Mexican American Legal Defense and Education Fund (MALDEF)
National Association for the Advancement of Colored People (NAACP)
National Conference of Puerto Rican Women
National Organization for Women (NOW)
National Political Congress of Black Women, Shirley Chisholm, Chair
National Society of Black Engineers
National Women's Law Center
New York Public Interest Research Group (NYPIRG)
Project on Equal Education Rights (PEER)
Sociologists for Women in Society
Women's Equity Action League (WEAL)
U.S. Delegate Walter E. Fauntroy (D.-DC)
U.S. Representative Claudine Schneider (R.-RI)
Eleanor Smeal

FairTest's Sarah Stockwell and NYPIRG Legislative Director Blair Horner at the FairTest Conference.

Input Sought for New York Test Bias Study

New York State's first-in-the-nation Advisory Committee on Fairness and Equity of Standardized Admissions Tests has launched its investigation of bias in the SAT and other university admissions tests. At its first full meeting, consulting psychologist Richard Barrett was elected Chairman and Cornell University's Jason Millman Vice-Chairman (see *Examiner* Vol. 2, No.2, for full list of members).

The New York Advisory Committee is seeking public input, including citations of "studies of test bias in post-secondary admissions or of alternative methods to monitor or reduce such bias," as well as what would be "important data to collect or studies to conduct to learn whether additional legislation is required for these standardized admissions tests".

Later this fall the Advisory Committee will issue a Request For Proposals seeking a consultant to conduct an analysis of item statistics provided by testing companies and to research related topics concerning the causes and solutions to test bias.

* * *

To get a copy of the RFP or give the Advisory Committee your suggestions, contact Chairman Richard Barrett, Ph.D.,
5 Riverview Place, Hastings-on-Hudson,
New York 10706.

Hypocrisy at Work

Despite their pronouncements that coaching for the SAT and PSAT doesn't affect scores, ETS and the College Board have developed a "comprehensive test-preparation curriculum for teachers, tutors and guidance counselors to use in preparing Hispanic students for the PSAT/NMSQT." The curriculum includes 22 lesson plans, which cover "all types of verbal and mathematical questions." An audiotape, a filmstrip on how to take the exam and a practice test are also part of the package.

Not surprisingly, ETS and CEEB researchers have found that the coaching kit and "associated activities" were "very successful" in increasing student confidence about taking the test. However, the study leaves the obvious question unanswered: did the test scores increase for students involved in the new program?

University Admissions Testing

New York Study Proves SAT Worthless

Data uncovered by the New York Public Interest Research Group (NYPIRG) demonstrates that the SAT is virtually meaningless in predicting first year grades for college students in New York State. In a new report, *Rolling "Loaded" Dice: An Analysis of SAT Use for Higher Education Admissions in New York State*, which analyzes information collected under New York's landmark Truth-in-Testing law, NYPIRG found that the SAT improved the average college's admissions decisions by just 6% above pure chance.

High school performance (grades or class rank) proved a much better indicator of college ability, making admissions decisions nearly 20 percent more accurate. The SAT's sole claim to utility is its supposed ability to predict first year college grades. Based on the study, NYPIRG concluded that "the SAT simply doesn't work, unless you are black, poor or female—and then it works against you." (NYPIRG Publications, 9 Murray Street, 3rd floor, New York, NY 10007; $5.00)

Gender-Biased Questions Found on the SAT

A new study has identified specific SAT questions which produce large gaps between male and female scores. Consistent with the test's general biases, many more questions favored boys: of 17 items favoring one sex by about 10% or more, 13 favored boys and only 4 favored girls, suggesting that ETS' "review processes" are inadequate to detect unfair bias. Researchers Phyllis Rosser (author of FairTest's report, *Sex-Bias in Testing: Why Women Lose Out*), Jim Loewen (Professor of Sociology at the University of Vermont), and Princeton Review coaching company President John Katzman administered an SAT to 1112 predominantly white and upper-middle class coaching students.

The 13 questions favoring boys included several about chores at a boys camp and one requiring the test taker to project a basketball team's winning record. This question favored boys by an astounding 27% (see sample questions). The result of such bias is that the SAT used in the study under-predicted girls' high school performance, as it does nationwide: girls in the sample reported better grades than boys in both English and math, yet scored about the same on the SAT-Verbal and lower on the SAT-Math.

SAT QUESTIONS WITH A LARGE MALE/FEMALE SCORE GAP

FAVORS MALES BY 27%	A high school basketball team has won 40% of its first 15 games. Beginning with the sixteenth game, how many games in a row does the team now have to win in order to have a 55 percent winning record?
FAVORS MALES BY 16%	MERCENARY:SOLDIER:: (A)censor:author *(B) hack:writer (C) agent:performer (D) fraud:artist (E) critic:subject
FAVORS FEMALES BY 10%	PENDANT:JEWELRY:: (A) frame:picture (B) cue:drama (C) violin:music *(D) mobile:sculpture (E) poetry:prose

* wanted answer

ACT Under Fire in Alabama

The Alabama Department of Education is being sued for misuse of American College Testing Assessment (ACT) scores in admitting students to the state's colleges of education. Currently applicants must score at least 16 on the ACT to be considered for teacher training programs. The plaintiff's lawyer, Terry Davis, claims the arbitrary use of ACT cutoff scores is racially discriminatory.

Davis had previously sued the Alabama Department of Education for using ACT scores as the sole factor in determining winners of the Alabama state scholarship program, a procedure that resulted in no Black students winning scholarships in 1982-83 and 1983-84. The court dismissed that case, ruling that the plaintiff, Alabama State University, had no standing. Soon thereafter, Alabama instituted a new procedure that is only marginally better for Black students. For more information, contact Attorney Terry Davis, (205) 264-2230.

(More University, p. 13.)

Grade School Testing

NAEP Report Finds No Math Improvement

Despite gains in basic arithmetic skills, particularly among minorities, higher order math skills and problem solving ability have not improved among students, according to *The Mathematics Report Card: Are We Measuring Up?* Only half the 17-year-olds were able to perform at or above the level of "moderately complex procedures and reasoning," and only 6.4% of 17-year-olds attained the highest measured level, "multi-step problem solving and algebra."

The report, issued by Educational Testing Service for the National Assessment of Educational Progress (NAEP), was based on a sample of 9, 13 and 17-year old students across the U.S.

To overcome the lack of higher order skills, the report urged fundamental revision in math instruction. It emphasized the need for teaching problem-solving and reasoning skills, and utilizing projects and group activities.

While the report concluded, "Students may spend far too much time on drill and practice," it did not address the effects of standardized testing on math curriculum and teaching. Yet a focus on raising test scores encourages rote drill and practice, precisely the methods of instruction that discourage the development of mathematical thinking and application.

The recently released NAEP report on reading also argued against the over-emphasis on basic skills (*Examiner* Vol. 2, No. 2). The two NAEP reports confirm the conclusion of FairTest's study, *Fallout from the Testing Explosion*, that "by narrowing the curriculum, standardized tests are undermining many of the most important goals of the current school improvement movement." (*Mathematics* can be ordered from NAEP, Rosedale Rd., Princeton, NJ 08541; $14.)

K-12 News In Brief

North Carolina, which last year ended funding of state-mandated tests for first and second graders, has now barred local school boards from using any standardized exams in those grades. A new law says, "Local school administrative units may use ['developmentally appropriate individualized assessment instruments'] provided to them by the State Board for first and second grade students, and shall not use standardized tests." State funds for the alternative assessments cannot be diverted to purchasing tests.

Mississippi's school superintendent has decided to discontinue the state-mandated kindergarten test, effective with the 1989-90 academic year. Mississippi teachers had protested use of the exam because it was driving curriculum away from the state's progressive, developmentally-oriented kindergarten program.

Test scores in **Clark County, Nevada** dropped in reading, math, and language, leading testing coordinator Judith Costa to proclaim, "I'm pleased." A school board study had concluded that test scores were inflated because of "too much emphasis on test preparation." The district decided to de-emphasize the scores by not using the tests to evaluate teachers and testing earlier in the school year (*Las Vegas Review-Journal*, 1/18/88).

Connecticut will become the first state in the nation to distribute financial aid to schools on the basis of test score results. Aid will be awarded both for districts with a high number of low-scoring students and to districts that improve their test scores (*Chronicle of Higher Education*, 5/25/88).

South Carolina will now offer a certificate, not a diploma, to students who complete high school course requirements but who do not pass the state's exit examination. The test will be given to tenth graders starting in 1989-90. Those who do not pass will be able to take it again in school programs until age 21, and through adult education programs thereafter (*Education Week*, 3/2/88).

A proposal for a similar gatekeeper exam in **New Jersey** has angered the New Jersey Education Association. In opposing an 11th grade test for graduation, the NJEA observed, "We don't believe that one instrument should determine whether a student graduates or not" (*Education Week*, 5/11/88).

Publications

The National Coalition of Advocates for Students (NCAS) has released *New Voices: Immigrant Students in U.S. Public Schools*. This comprehensive study includes a sharp criticism of the "lockstep use of standardized tests" The report suggests radically changed intake and assessment procedures and a wider restructuring of schools. (*New Voices*, NCAS, 100 Boylston St., Boston, MA 02116; $11.95.)

A Conspiracy of Good Intentions: America's Textbook Fiasco, by Harriet Tyson-Bernstein of the Council for Basic Education, claims that textbook quality has declined to the point of unreadability. One cause is configuring texts to standardized tests. CBE urges state policymakers to separate testing programs from textbook selection. (CBE, 725 15th St., N.W., Washington, DC 20005; $13.)

Eating Breakfast Raises Test Scores

Students who ate breakfast made greater gains on test scores than students who did not, according to a study reported to the Ambulatory Pediatric Association. This evidence is aiding a push to expand the federal breakfast program. About 3.6 million students receive subsidized breakfasts, compared with the 24 million who receive subsidized lunches (*Education Week*, 6/8/88).

Grade School Testing

Arizona Reduces School Tests

The state of Arizona no longer requires local school districts to administer standardized tests to first and twelfth grade students. A grass roots parent and teacher campaign persuaded the legislature to end the testing mandate, explained Ann Gould, a teacher and co-chair of the Committee for Effective School Evaluation (CESE), which led the effort.

Districts can opt to continue first and twelfth grade testing and request the state to pay for it, and the state will test a sample of students in the two grades. Mandated exams continue in grades 2-11. The law also called for research on statewide testing and alternatives.

According to Gould, the biggest obstacle CESE faced was "the fear of losing accountability." Teachers who opposed extensive testing were criticized for avoiding responsibility for their students' performance. Opponents claimed CESE's initial proposal to eliminate testing in both grades 1 and 2 would leave the state with no information about the quality of education before third grade.

Efforts by CESE to discredit the tests, in part for their inaccuracy for Native American and Hispanic students, met the response that they were "the best we have." Victory came through the argument that the test placed too much pressure on young children. The 12th-grade test was discredited because exam results were not reported until after students graduated. Thus, no one took the test seriously.

CESE hopes to continue to reduce mandated testing in Arizona and calls for using standardized exams "within reason" rather than treating the scores as "magic numbers."

"The whole school year is geared to preparing for the April test," explained Gould. Test scores are "front page news" and high scores in some communities are touted by real estate agents. Battles are now shaping up over whether districts will continue to test in grades 1 and 12. Gould explained that some districts still require kindergarten tests, though the state eliminated that mandate a few years ago.

In addition to working within districts, CESE wants to continue to educate legislative leaders. Some members of a joint committee of educators and legislators have already concluded that the educational goals they seek in Arizona are not well measured by standardized, norm-referenced tests. A state financial crunch is aiding calls to test every other grade rather than every grade.

Ann Cook of Community Studies, New York, presenting the workshop on K-12 testing.

Segregation by Testing

The Mississippi chapter of the NAACP has asked the U.S. Justice Department to file suit against the state's high school exit examination on the grounds that the test discriminates against Blacks. The NAACP's Morris Kinsey termed the test "totally unconstitutional and totally unfair to Black students."

The Black failure rate on the Functional Literacy Exam has been three times that of the White. The state claims that the score discrepancy on the exit test "is a poverty indicator."

But Kinsey says, "Southern educators and administrators alike have used test scores to keep segregation. It's 1988 and we still have the same racially segregated school systems and classes—in the name of testing."

Kinsey blamed inferior education for the score differential and noted that Blacks are disproportionately enrolled in remedial and vocational classes that do not prepare students for the test.

Testing for academic placement is common in Mississippi. From the advent of the 1972 desegregation order until recently, the state used the California Achievement Test (CAT). The state now mandates the Stanford for grades K, 1, 4 and 6 (though the superintendent will drop the kindergarten test next year). Children are grouped by test scores with most Blacks in the lower levels.

Though 58% of the state's half-million public school students are Black, about 85% of the students in special education and Title 1 (federally funded remedial programs) are Black. As a result, over half the Black students are in these two categories.

Nathaniel Douglas of the Justice Department informed FairTest that the Department is investigating the situation.

FairTest Conference

Senator Gore's Keynote Address

...It is especially distressing to see young children subjected to the bias inherent in some tests. Many states use standardized tests to determine the kindergarten eligibility of five and six-year olds. In one state, 40% of the black five-year olds in that state failed a "Ready-or-Not" exam— and as a result, were placed in remedial classrooms before they even began formal schooling. Such a serious decision about any young child's future ought to be based on more than a single test.

Testing has become more prevalent as parents and administrators around the country insist on quantifiable results in education. Many states have adopted high school competency tests, in an understandable attempt to set minimum academic standards....But nationwide, Black, Hispanic and low-income students fail these exams at unusually high rates.

This discrepancy also turns up in college testing....Despite indications of bias in these tests, over 750 government and private scholarship agencies rely heavily on the test scores of the SAT or the ACT to help determine recipients. Unfortunately, this means that women, Blacks, Hispanics and other minorities tend to be underrepresented as scholarship finalists.

To some extent, these results reflect existing social inequities. For example, a poor child may test poorly because her parents could not afford to give her a decent breakfast. But the testing bias seems more deeply ingrained than that. And with so many educational and economic opportunities turning on test results, standardized tests may serve to make those existing inequities worse....

In the last few years, the "Truth-in-Testing" movement has helped to overcome some bias in higher education admissions tests. But testmakers still bear a special responsibility to develop unbiased tests that reflect the cultural diversity in our society....

...Our progress [as a nation] won't be complete if education and professions are closed to talented individuals because biased exams prevent them from competing on a level playing field....Our goal in education and in testing must be the same as our goal for our nation: that all individuals, regardless of sex, race, religion, ethnic origin, or physical disability shall have an equal chance to achieve.

Breaking Down the Barriers

(from page 1)

Alabama Attorney Don Watkins

cessful legislative campaign barring state-funded testing in grades one and two in North Carolina; and Truth-in- Testing and scholarship reforms in New York State (see this and previous *Examiners* for details). Besides chronicling their success, each of the speakers addressed the spirit needed to continue the battle.

"We have a saying in Alabama," said attorney Don Watkins of the victory against teacher testing. "Don't let the same snake bite you twice. We have killed the snake that had us in 1981, and we do not intend to let that snake or any snake comparable in nature bite us again."

Page McCullough of the Atlantic Center for Research in Education (ACRE) added, "It is possible to stop some of the current testing craziness and the abuses that it is heaping on little kids and their teachers. If you can do it in North Carolina, you can do it anywhere." And Blair Horner of the New York Public Interest Research Group (NYPIRG) concluded, "This is not as much an education issue as it is a political issue. When you approach this as consumers organizing against the manufacturer of a faulty product, you have a much better concept of how to proceed and how to win."

Many of the workshops which followed focused on how to replicate and extend these victories. Signs of progress are visible in each area of testing though many new problems loom on the horizon.

Grade School Testing

Participants reported that educators, policy analysts and political figures around the nation are beginning to reassess their commitment to testing. The issues of bias, reliability and validity, and the role of test scores in shaping curriculum and tracking students, have all contributed to a growing awareness of the dangers of the overuse and misuse of standardized, multiple-choice exams in the schools. Yet, progress has been uneven: while North Carolina and Arizona have voted to reduce mandatory testing, Georgia initiated testing for entrance to first grade.

Recent legislative victories along with the new positions of the National Association for the Education of Young Children and the National Association of Early-Childhood Specialists opposing testing of young children, led to wide agreement that campaigns must be developed to halt the mass testing of children before grade 3.

There was also consensus about the need to expose and combat the use of testing to enforce or justify segregation by race and class. Federal legislation or regulations to bring all testing under guidelines such as those in PL 94-142 (special education) were proposed to help

FairTest Conference

Closing Remarks by Vito Perrone

....Much more needs to be done. And what needs doing goes, of course, beyond tests.

This isn't to say that we should not continue to advocate less use of tests for critical decisions about individuals, that we shouldn't continue to challenge the assumptions of tests, that we should withdraw from our critique of the full technology of testing or not make clear the many negative consequences of tests. All of that work obviously needs to go forward.

But alongside it, there needs to be more attention to the context of tests, greater understanding of why testing remains so popular, more serious pursuits of alternative ways of responding to the legitimate interests in accountability and real knowledge. The politics, in particular, need more attention and understanding.

In spite of us, the belief that tests provide important information and are, or could be, beneficial to children, teachers, and employers is widespread....Such an understanding is connected to a particular vision of schools, places for children to fit and not places to fit children. That vision, unfortunately, has widespread support.

When schools and the purposes of schools are defined narrowly in terms of particular, discrete skills and fairly precise bodies of information, tests of the kind we know appear on the surface to be logical, even useful. On the other hand, when schools and their purposes are defined more broadly, and I would posit, therefore more powerfully, thought about in terms of dispositions, attitudes, understandings, and knowledge, with skills being integral to and not apart from them, the tests we know increasingly appear less useful, even crude, and more thoughtful alternatives quite naturally emerge....

All of this is to suggest that we need more than is now the case to present more articulately a powerful vision of schools that is fully democratic, serving all children and young people, and rooted in the life of important communities. We need to link more firmly *powerful* schools and *powerful* assessment as we go forward with our advocacies. They need to be discussed together, as part of the same fabric. That same argument could be offered for making more clear what it means to live and work in a democratic society. It is out of that linkage—and possibly only from that linkage—that alternatives of consequence can emerge....

So, where do we go from here? We continue to do what we have done. We challenge inequity wherever it appears. We create as much discomfort as possible. We inform. We make use of the media, the courts and the legislative halls. And we continue describing ever more clearly, as well as supporting those who are implementing those descriptions, the fullest meaning of a powerful democratic education and the related and necessary processes of evaluation.

---

halt the segregative uses of testing.

The success of truth-in-testing efforts around the SAT, coupled with questions about bias, validity and reliability of other tests, led to support for applying similar legislation to school tests. Materials about tests and their effects must also be developed and distributed in ways that are accessible and useful to community, parent, and teacher groups, as well as administrators, legislators and policy personnel. Training sessions could be a useful adjunct to resource development.

University Admissions

Conference participants strongly supported efforts to stop misuse of PSAT scores by the National Merit Scholarship Corporation (see story this issue). Success in this campaign would simultaneously provide substantial tuition assistance and other benefits to female and minority students and continue to expose the bias and limited predictive value of the SAT.

Other issues generating substantial interest included how to use the coaching controversy to force further reforms, and the causes of racial and gender bias in admissions tests. Cut off scores continue to plague those concerned with opening up the college admissions process, despite the fact that test-makers admit this is a misuse of test results. Since about 40 percent of all colleges set minimum SAT or ACT scores for admission, the gender and racial score gaps ensure that minorities and women will remain at a distinct disadvantage in attempting to obtain higher education.

Among long term strategies discussed was the possibility of convincing school accrediting agencies to stop considering test scores in their accreditation decisions. Legal actions to stop the misuse of test scores are also a possibility.

Employment Tests

Teacher licensing tests remained a top priority since use of the exams is sharply reducing the proportion of minority teachers at the same time that the Black and Hispanic student population is growing rapidly. Despite successful litigation in Alabama and Georgia, use of the tests continues to spread.

Vito Perrone of the Carnegie Foundation for the Advancement of Teaching.

The absence of predictive validity for any teacher test, along with their pro-

(continued on page 10)

Breaking Down the Barriers

(from page 9)

nounced racial impact, may make some exams vulnerable to litigation. However, success against tests developed by National Evaluation Systems has been counter-balanced by repeated failures in legal challenges against the NTE. Coaching and aligning college curricula with the test may help reduce the negative impact on minorities, but participants were concerned that such techniques detract from time spent training teachers in useful skills and knowledge. Thus, legal, legislative, and public pressure must be brought to bear to halt use of biased and inaccurate tests.

Similar situations apply with regard to most other occupational licensing exams as demonstrated in a recent Wisconsin study of differential pass rates (see p. 11). More information needs to be exchanged among groups working to replace unfair and irrelevant licensing requirements. State and local civil service tests also were suggested as a potential area for activists since state equal opportunity laws may offer avenues for intervention.

Moving Ahead

For those in attendance, "Breaking Down the Barriers" was an energizing experience. The rich knowledge and experience of the conference participants expanded the testing reform movement's vision and strengthened its hope.

But much more remains to be done. As Carnegie Foundation for the Advancement of Teaching President Vito Perrone remarked in an eloquent conference closing address, "There needs to be more attention to the context of tests, greater understanding of why testing remains so popular, more serious pursuits of alternative ways of responding to the legitimate interests in accountability and real knowledge." That is FairTest's challenge for the coming years.

Quote:

"Testing is like sex. It was a lot healthier when it was more private."

—Donald Stewart, President, College Board (*Education Daily*, 3/11/88).

(l- r) FairTest Managing Director Monty Neill, Harold Dent of the Association of Black Psychologists, and David Feldman of Tufts University at FairTest Conference

Allen: For or Against IQ Tests?

Conference sessions were often enthusiastic and occasionally contentious. For example, Dr. Harold Dent of the Association of Black Psychologists, in discussing the successful campaign in California to bar the use of IQ tests in placing Black children in special education classes (see *Examiner* Vol. 2, No. 1), criticized the U.S. Civil Rights Commission for seeking to circumvent the *Larry P.* decision. Commissioner William Allen replied from the audience that changing the court ruling was his personal crusade, not an official Commission position.

After further debate, Allen admitted that the problem with IQ tests was that they lacked validity for anyone. Dent then pointed out that they agreed on this point, but that Allen was trying to reinstate IQ tests for Blacks, not bar them for everyone. Other participants joined Dent in asking Allen to aid a drive to stop all placement of children based on IQ tests.

Conference Tapes

Audio cassette tapes of the FairTest conference can be ordered from TCSD Sight and Sound Productions, 6006 Greenbelt Road, Suite 165, Greenbelt, Maryland 20770. Cost is $7/tape plus $1.50 postage for first 3, $3 for 4 or more tapes.

The following plenaries and workshops are available:

Opening Session (1 tape)

Morning large group sessions (3):
Grade School Testing (*morning*)
University Admissions Testing
Employment Testing

Afternoon Workshops (12):
Grade School Testing (*afternoon*)
Linguistic and Cultural Bias in Standardized Tests
Discrimination in Test-based Scholarship Awards
Problems of Teacher Testing
The National Assessment of Educational Progress and National Tests
Employment Tests: Can They Be Fair and Relevant?
Organizing a Testing Reform Campaign
Alternatives to Testing Schoolchildren
Exploring the Gender Gap on the SAT
Litigating Teacher Tests
Using the Coaching Controversy to Promote Test Reforms
Auditing High Stakes Testing

Closing Session (1)
(Vito Perrone)

Employment Testing

Wisconsin Examines Bias in Licensing Tests

The Wisconsin Department of Regulation and Licensing has uncovered a pattern of adverse impact against minorities in the exams that the state requires for entrance into more than half a dozen occupations. Pass rates for Blacks, Hispanics, and other minorities were found to lag far behind Whites on tests for accountants, cosmetologists, engineers, nurses, private detectives, and real estate brokers.

An analysis of state data, conducted by the Center for Public Representation, suggests that legal action under the Wisconsin Fair Employment Act might force an overhaul of testing requirements. Wisconsin officials have hired a researcher to study why minorities perform so poorly on the cosmetology and nursing exams, tests which Blacks are twice as likely to fail as Whites.

Copies of the report, *Regulating Occupations: Legal Challenges to Licensing Examinations in Wisconsin*, which includes data on differential pass rates and an analysis of legal options under state and federal statutes, are available for $5.00 prepaid ($2.00 extra fee for invoicing) from the Center for Public Representation, 520 University Avenue, Madison, WI 53703.

A working session at the FairTest Conference.

Coaching Works for Jobs Tests, Too

Convinced that well-designed test preparation courses can significantly increase performance on college admissions exams, researchers have begun to investigate whether coaching can have a similar impact on occupational licensing test scores.

In one pilot program, University of North Carolina School of Medicine researcher Dr. Henry Frierson ran test prep experiments with minority medical and nursing school students. Reporting on the results to the American Education Research Association, Frierson concluded that coaching clearly works: those who received test-taking instruction had far higher passing rates on required licensing exams than their peers who were assigned to a control group. Better results were achieved through a combination of coaching classes and "student driven" learning teams (TTI/LT) than through coaching alone.

Based on these results, Frierson recommends similar programs for all schools interested in increasing minority retention and pass rates on licensing exams.

A copy of the paper can be obtained from Dr. Henry Frierson, Office of Research and Development, School Of Medicine, University of North Carolina, Chapel Hill, NC 27599-7530.

(More employment, p. 13)

NTE Use Wipes Out Maryland Minority Teachers

The growing controversy surrounding the use of the NTE (formerly the National Teachers Examination) as a licensing requirement for classroom educators has reached the suburbs of Washington, D.C. In Prince George's County, Maryland, nearly half of the 110 "provisional" teachers required to take the test to keep their jobs did not pass. Most of those who failed were minority group members hired less than one year ago under a plan to overhaul the local school system.

The county terminated all those who failed the test even though, as a school district official admitted, "Many of these teachers were evaluated very highly, they proved to be very good reliable teachers." The county has decided to force all new applicants to pass the test, before initial hiring. Teachers union officials have criticized the NTE requirement, but no legal challenge has yet been filed.

The impact of the NTE in Maryland should not be surprising. A soon-to-be published study by University of North Florida Prof. G. Pritchy Smith concludes that 37,700 minority teaching candidates, including 21,500 Blacks and 10,000 Hispanics, have been kept out of the teaching profession by the use of similar licensing tests around the nation. Yet, no study has ever proven the NTE has any validity for predicting whether an applicant will be a good teacher.

Even the Educational Testing Service, the NTE's manufacturer, agrees the exam must be replaced. The company recently announced plans to create a successor test by 1992. Unfortunately, teaching candidates in Maryland and some 20 other states will still be required to pass the flawed NTE for the next four years.

State-Mandated Tests, 1986-87

STATE	NUMBER OF TESTS	SCHOOL ENROLLMENT	STATE-MANDATED TESTS PER STUDENT
Alabama	720,000	733,735	0.98
Alaska	0	107,973	0.00
Arizona	480,000	534,538	0.90
Arkansas	179,000	437,438	0.41
California	1,420,000	4,377,989	0.32
Colorado	208,800	558,415	0.37
Connecticut	99,000	468,847	0.21
Delaware	60,000	94,410	0.64
District of Columbia	119,000	85,612	1.39
Florida	437,352	1,607,320	0.27
Georgia	1,020,000	1,096,425	0.93
Hawaii	75,000	164,640	0.46
Idaho	30,000	208,391	0.14
Illinois	418,000	1,825,185	0.23
Indiana	211,000	966,780	0.22
Iowa	0	481,286	0.00
Kansas	135,000	416,091	0.32
Kentucky	650,000	642,778	1.01
Louisiana	390,000	795,188	0.49
Maine	48,000	211,752	0.23
Maryland	168,000	675,747	0.25
Massachusetts	410,000	833,918	0.49
Michigan	326,285	1,681,880	0.19
Minnesota	0	711,134	0.00
Mississippi	240,000	498,639	0.48
Missouri	664,000	800,606	0.83
Montana	0	153,327	0.00
Nebraska	20,508	267,139	0.08
Nevada	55,000	161,239	0.34
New Hampshire	30,000	163,717	0.18
New Jersey	630,000	1,107,467	0.57
New Mexico	80,000	281,943	0.28
New York	1,691,000	2,607,719	0.65
North Carolina	1,085,000	1,085,248	1.00
North Dakota	0	118,703	0.00
Ohio	0	1,793,508	0.00
Oklahoma	234,000	593,183	0.39
Oregon	15,000	449,307	0.03
Pennsylvania	538,212	1,674,161	0.32
Rhode Island	40,000	134,126	0.30
South Carolina	450,000	611,629	0.74
South Dakota	50,000	125,458	0.40
Tennessee	500,000	818,073	0.61
Texas	1,500,000	3,209,515	0.47
Utah	258,000	415,994	0.62
Vermont	0	92,112	0.00
Virginia	1,069,000	975,135	1.10
Washington	182,000	761,428	0.24
West Virginia	200,000	351,837	0.57
Wisconsin	309,000	767,819	0.40
Wyoming	0	100,955	0.00
UNITED STATES	17,545,159	39,837,459	0.44

Testing Explosion

(from page 16)

should measure pertinent, not extraneous, knowledge differences, and should not be biased against any groups in the population.

- Tests should be open to test-takers, their parents, and independent researchers so that questions and answers can be examined and test-makers' claims can be subject to independent scrutiny. The testing industry must be brought under public control.
- Test results must be properly interpreted and employed. They must not be used as the sole or primary criterion in "high stakes" decisions, such as grade promotion, school graduation or entrance into special programs.
- New assessment instruments must be developed as alternatives and supplements to standardized multiple-choice tests.

* * *

Fallout From the Testing Explosion is available for $8.95. Use the order form on p.15 to place your order today. And don't forget to get extra copies for local school officials and policy makers.

New York Court Upholds Band Scoring

In a blow to those who believe small differences in test scores have a profound meaning, the New York State Court of Appeals has unanimously approved the use of "zone" or "band" scoring in Civil Service tests. Under the reformed system, all applicants who score within a given range may be grouped together. By assigning jobs randomly to those within the same zone, the percentage of minorities offered employment can dramatically increase.

The state's Civil Service Employees Association had charged that use of zone scoring would dilute "merit" and "fitness". But the court held that the differences in knowledge or ability between candidates in any band are likely to be marginal. Thus, zone scoring may be used in New York any time a government agency shows that a job-related quality cannot be accurately measured by a test.

Some employee groups may try to block zone scoring on a case-by-case basis. But, for now, zone scoring is permissible throughout the Empire State.

Feds Set Pace by Abandoning Biased Tests

In a surprising move, the Reagan Administration has decided to abandon its quest for a new selection test for U.S. government workers. For six years, the federal Office of Personnel Management (OPM) has tried to design a replacement for the Professional and Administrative Career Examination (PACE). Before the Carter Administration settled an anti-discrimination law suit by stopping the use of PACE, up to ninety-five percent of black applicants had been screened out by the test.

Beginning next April, OPM proposes that anyone with a college grade point average above a specified level (3.00 to 3.25 on a 4.00 point scale) could automatically be offered employment with no required testing. Others will be evaluated based on their performance on "jobs related skills tests" and on a new Individual Achievement Record. Government personnel officials say the reform will make federal employment more attractive and encourage stepped up recruitment at historically-Black institutions.

Civil rights advocates and test reformers are heartened by the decision, but plan to closely monitor implementation of the new system.

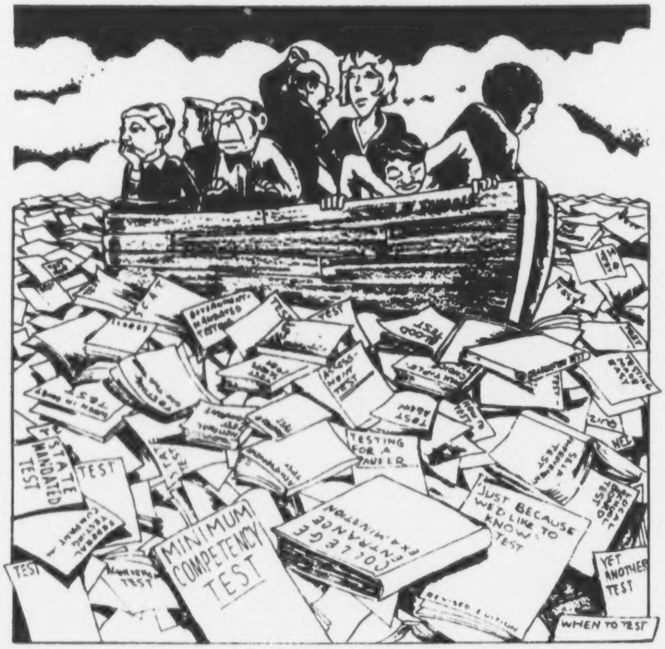


Illustration by Ned Shaw

The Case Against the SAT

James Crouse and Dale Trusheim, in their book on the SAT and its uses by colleges and universities, argue that "*despite* its ability to predict educational success, the SAT is unnecessary," (emphasis in original). The authors demonstrate that the SAT has an adverse impact on Black and low-income applicants and conclude that the SAT neither helps colleges make better admissions decisions nor helps applicants select colleges. (*The Case Against the SAT*, The University of Chicago Press, 1988, 224pp. $22.50.)

None of the Above

FairTest is now able to offer David Owen's book, *None of the Above: Behind the Myth of Scholastic Aptitude*, at a substantial discount. Use the order form on page 15 to get your copy now.

THE NEW YORK TIMES **EDUCATION** WEDNESDAY, JUNE 29, 1988

Groups Charge Bias in Merit Scholarship Testing

By LEE A. DANIELS

The competition to determine winners of the prestigious National Merit Scholarship awards for high school students is substantially flawed by reliance on a test that is "biased on the basis of sex, race and family income," according to an group that monitors standardized testing.

The group, the National Center for Fair and Open Testing, has banded together with several legal, feminist, civil rights and other educational advocacy groups to urge changes in the test program. In a letter released yesterday they asked the National Merit Scholarship Corporation, which administers the program, to use other criteria, such as high school grades and class rank, to determine who receives the scholarships.

National Merit, based in Evanston, Ill., awards nearly 6,000 scholarships worth more than $23 million a year.

An Academic Focus

Officials of the scholarship corporation defended the use of a standardized test to select scholarship recipients, saying they had found no flaws. The officials also said that in fact the test questions were now only one of several factors they considered in chosing National Merit scholars.

"Ours is an open competition that is very keenly academically focused," said Marianne Roderick, senior vice present of the corporation.

In the letter to the scholarship corporation, the National Center for Fair and Open Testing, also called FairTest, said, "We are alarmed and dismayed by the relatively small proportion of minority and female winners of the scholarships you administer."

Different Experiences Cited

FairTest said it had estimated that nearly two-thirds of the National Merit awards went to males even though female students as a group earned higher grades in high school and college.

The organization also estimated that members of minority groups, who make up more than a fourth of high school seniors, received a much smaller share of scholarships. This is the case, FairTest said, even though about 11 percent of National Merit Scholarships are earmarked for black and Hispanic students through the Achievement Program and the National Hispanic Scholars Program, respectively.

Sample Questions

Word association questions taken from recent Scholastic Aptitude Tests. Groups calling for changes in tests say such questions discriminate against students who are unfamiliar with activities of upper-middle-class Americans. Asterisk denotes answer sought by testing service.

RACQUET:TENNIS
springboard:diver
horse:polo
glove:boxing
club:golf*
gun:hunting

HEIRLOOM:INHERITANCE
payment:currency
belongings:receipt
land:construction
legacy:bill
booty:plunder*

CONSERVATORY:MUSIC
anthology:books
aerie:birds
bivouac:army
seminary:religion*
arbor:grapes

BRIDLE:HORSE
bone:dog
olive branch:dove
valor:soldier
precept:conduct*
devotion:duty

Source: National Center for Fair and Open Testing

Ms. Roderick replied that the differences in test performances among women, blacks, Hispanics and white males "doesn't mean that the test itself is biased." But. she said, "it does indicate that differences exist in the academic experiences of those groups."

Use Called Excessive

The newest charges by FairTest and the steadfast defense by testing services underscore how contentious the debate over the validity and use of standardized tests remains.

Last year, FairTest charged that another examination, the Scholastic Aptitude Test, which most colleges use as one measure of which high school students they admit, was biased against girls, again pointing out that they as a group scored lower on that test even though they earned better grades than boys.

And last month, the Cambridge, Mass.,-based group released a report that said there was an excessive use of standardized tests by elementary and secondary school educators.

The National Merit scholarships are among the best-known and most sought-after grants for college. Ms. Roderick said that this year 1.2 million students in 19,000 secondary schools took the qualifying test for the competition.

Winnowing Process

The scholarship corporation then winnows that pool first to 15,000 semifinalists, then to 14,000 finalists, and, finally, to 6,000 scholarship winners. Last year the corporation awarded more than $23.5 million in scholarships, excluding the programs for minorities, Ms. Roderick said.

Among the groups signing the FairTest letter were the National Association for the Advancement of Colored People, the Mexican American Legal Defense and Education Fund, the Asian American Legal Defense and Education Fund and the National Organization for Women.

The focus of the latest contentions by FairTest is the Preliminary Scholastic Aptitude Test/National Merit Scholarship Qualifying Test, which students take in October of their junior year.

References to Tennis and Polo

FairTest contended that the disparities it found in the proportion of scholarships won by black, Hispanic, and female students resulted from a bias inherent in the examination and based on sex, race, and class. The letter included examples from what the group said were recent aptitude tests. FairTest said the sample questions showed that students who took the test were expected to be familiar with upper-middle-class sports like sailing, tennis and polo and dance terms like minuets and pirouettes.

The FairTest letter said that there were far more test questions based on what could be considered normal activities of white males, rather than of females or blacks or Hispanics. It said that many female and minority students might find such questions disconcerting.

FairTest also questioned whether the two-hour, multiple-choice examination on which the National Merit competition is based was the best way alone to evaluate the students. FairTest asserted that many students "who have worked hard and with distinction throughout four years of high school had their hopes dashed by this same exam."

Sarah Stockwell, a spokeswoman for FairTest, maintained that using the test in awarding scholarship money meant that those students most in need of college aid were least likely to get it.

Subscribe to FairTest Publications

FairTest publications will keep you up-to-date on the growing movement for fair & open testing. You will be able to do more than just read about FairTest's campaigns—you can help make them successful!

_____ *FairTest Examiner*, published quarterly...Individuals $15/year; Institutional checks $25

_____ *Fallout from the Testing Explosion: How 100 Million Standardized Exams Undermine Equity and Excellence in America's Public Schools*, by Noe Medina and Monty Neill with FairTest staff (45 pages)...$8.95

_____ *Standardized Testing Reform Sourcebook, 1988* (200 pages)...$25.
(Currently out of print, available Summer, 1988)

_____ *Sex Bias in College Admissions Tests: Why Women Lose Out*,
by Phyllis Rosser and FairTest staff, 2nd edition (40 pages)...$7.95

_____ *The Reign of ETS*, by Allan Nairn and Ralph Nader (550 pages)...$30

_____ *Beyond Standardized Tests: Admissions Alternatives that Work*,
by Amy Allina with FairTest staff (15 pages)...$5.50

_____ *None of the Above: Behind the Myth of Scholastic Aptitude*, by David Owen (327 pp. lists at $7.95)...$5.95

Updates, issued bi-annually (spring & fall). $7/year/topic:

_____ *Testing Reform Initiatives*
_____ *K-12th Grade Testing*
_____ *University Admissions Tests*
_____ *Teacher Testing*
_____ *Employment Testing*
_____ *Sex Bias in Testing*
_____ *Test Coaching*

Order total _____

Tax-deductible contribution _____

TOTAL ENCLOSED _____
(make check payable to FairTest)

_____ Back issues of *FairTest Examiner* $4 each: Vol___, No.___ (all 5 for $18.00)

_____**Moving.** I have moved/will be moving. My new address is (Please enclose old label or write old address.):

Name ______________________________ Title ______________________________
Organization ______________________________ Phone No. ______________________________
Address ______________________________ Zip ______________________________

Make checks payable to FairTest and mail to: FairTest, P.O. Box 1272, Harvard Square Station, Cambridge, MA 02238

Fallout from the Testing Explosion

More than 100 million standardized, multiple choice exams were administered to U.S. public school students during the 1986-87 school year, according to a new FairTest report, *Fallout from the Testing Explosion: How 100 Million Standardized Exams Undermine Equity and Excellence in America's Public Schools.*

The comprehensive study found that many tests are inaccurate, inconsistent or biased against minority, female and low-income students. Instead of promoting quality, accountability and equal opportunity, tests shift authority into the hands of an unregulated testing industry.

The overuse and misuse of testing also contributes to narrowing the curriculum. That, in turn, inhibits students' acquiring higher order skills. Over-reliance on the tests feeds into poor textbooks, grade retention and tracking policies that do not work, and a higher dropout rate.

Relying on standardized tests for decision making, the report concludes, "will lead to a *worse, not better*" educational system.

Too Many Exams

FairTest surveyed school officials in every state and the District of Columbia to discover how many tests are mandated. Nationally, states require the administration of over 17.5 million achievement, competency or basic skills tests or test batteries annually to nearly 40 million students.

On average, each public school student must take a state-mandated exam nearly every other year (see Table, p. 12). Yet this is only the tip of the testing iceberg.

To get the whole story, FairTest surveyed every large school district and a sample of medium and small districts on their use of locally-mandated standardized achievement tests. Calculating from the samples, FairTest concluded that nearly 39 million locally mandated exams were given, or one per student per year in the U.S. public schools.

Relying on other studies of the amount of testing done for Title 1 remedial programs, special education, college admissions and other purposes, the report concludes that at least 94-105 million standardized exams are administered yearly to public school students. And this figure does not include many IQ and "developmental" tests nor exams for bilingual students.

In general, larger districts test more than medium or small sized districts, and southern states test most frequently. All but one of the states which did not mandate testing had low minority populations.

A Tool for Reform

To make the report most useful to educators and policymakers, it includes a summary of critiques of testing such as the lack of validity, low reliability, and race, gender, class and cultural biases caused by constructing tests around the social realities of suburban, white children. The report also includes an annotated bibliography as a guide to further information about each of the problem areas.

Fallout from the Testing Explosion explains that heavy reliance on these instruments often has negative effects on pedagogy and curriculum, teacher behavior, and student achievement. It makes a strong case that most exams ought to have a label attached: *Caution, misuse of standardized tests can be dangerous to the nation's educational health.*

The report concludes with a number of reform recommendations. Among them are:

- Tests should be properly constructed, validated and administered. They

(continued on page 13)

FairTest
National Center for Fair & Open Testing
P.O. Box 1272
Harvard Square Station
Cambridge, MA 02238
(617) 864-4810

Address Correction Requested

Bulk Rate
U.S. Postage
PAID
Cambridge, MA
Permit No. 53278

22
Journal Printing Company
Revere, MA 02151

Complimentary Copy.
Please subscribe today.
See page 15.

FairTest

EXAMINER

The National Center for Fair & Open Testing Volume 2, Number 4 Fall 1988

Chorus for Change Grows Louder

National Merit Scholarship Corp. Digs In Its Heels

The National Merit Scholarship Corporation (NMSC) has begun to respond to the concerns of the nearly 50 civil rights, feminist and student organizations and leaders who have signed on to FairTest's letter criticizing the NMSC's reliance on a biased test as the sole qualifying factor for millions of dollars in scholarships.

Though National Merit leaders have dug in their heels in public, they are scrambling behind the scenes to justify NMSC policies.

In an uninformative letter to FairTest (and similar statements to the news media) NMSC President Elizabeth Jacka said, "A prolonged and time-consuming exchange and protracted debate...would be unproductive." Meanwhile, NMSC has sent a lengthy memo to its corporate funders trying to defend its current procedures. Not surprisingly, the memo was written by staff members of the Educational Testing Service, the company that manufactures the Preliminary Scholastic Aptitude Test/National Merit Scholarship Qualifying Test (PSAT/NMSQT).

The six-page memo demonstrates how NMSC and ETS persist in avoiding the main points raised in the public letter demanding changes in the scholarship selection system. Following are just a few examples of how NMSC and ETS continue to ignore and distort arguments to deflect attention from the real issues raised by FairTest.

- Though the issue raised by FairTest and its allies has always been the process for selecting *semi-finalists* (which defines the group eligible to compete for scholarships), the NMSC/ETS memo defends and explains in great detail the process for choosing *finalists* and *winners*. The fact is that reliance on the PSAT/NMSQT as the sole criterion for determining semifinalists guarantees that an unfairly small portion of scholarship winners will be female, and an even smaller proportion will be members of minority groups. Belaboring the point that factors other than test scores are weighed heavily in choosing winners, since "the test scores of all Finalists are extremely high," obscures the real issue. How could

(continued on page 8)

California Cheating Scandal

Thirty nine California school districts have admitted to cheating on the 1986 California Assessment Program (CAP) standardized tests. The cheating was done by teachers who altered answers on completed tests or encouraged students to change their responses.

The California Department of Public Instruction launched an investigation of 1985-86 CAP test results because of widespread allegations of cheating or inappropriate testing procedures.

The investigation was prompted when computer scoring machines discovered that 40 schools statewide had unusually high numbers of erasure marks on the computer scoring forms. All but one of the school districts have admitted to cheating on the CAP.

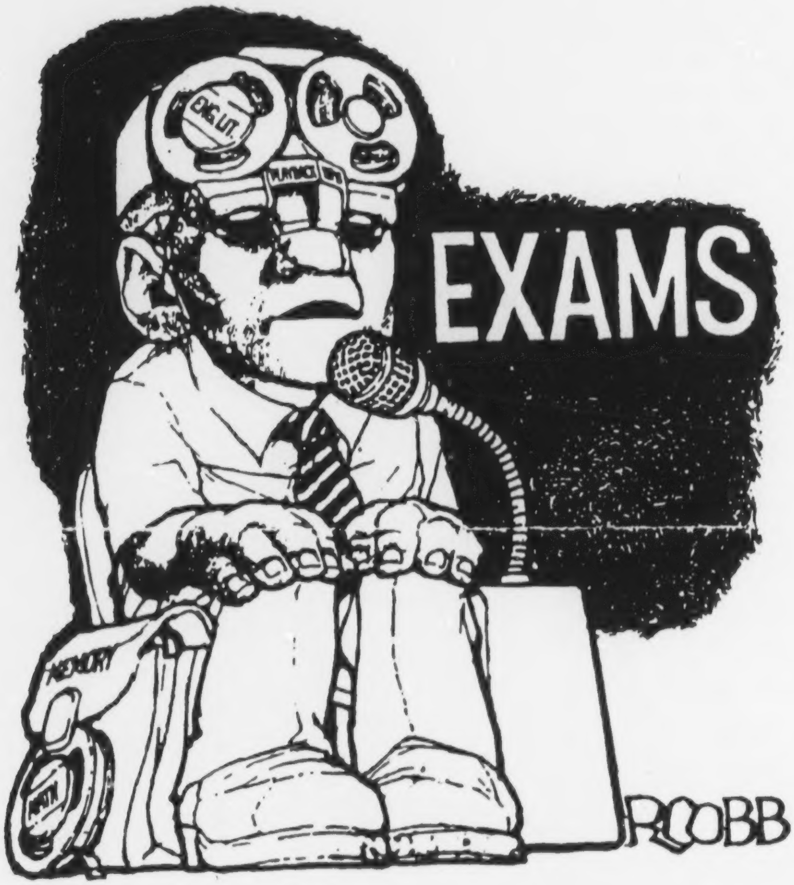
Because a number of local schools were involved, the Los Angeles Unified District has issued new guidelines to teachers for proper testing procedures and security. In addition, the district has hired FairTest Board member Ron Vera to help investigate the scandal.

California Superintendent of Public Instruction Bill Honig rejected the argument that extreme pressure on teachers to improve test scores invites cheating. But many teachers admitted

(continued on page 6)

CONTENTS

New Testing Code	2
Groups Oppose Testing	3
Coaching Babies	4
Gesell Challenged	5
New Chapter I Evaluations?	6
NMSC Letter Signers	8
NACAC Urges Proper Use	9
New York Reverses Reform	10
Prop. 48 and Black Athletes	10
Gifford Commission	12
Visiting Committee Hits ETS	13
NTE Change Announced	16



The National Center for Fair & Open Testing (FairTest)
P.O. Box 1272, Harvard Square Station
Cambridge, MA 02238; telephone (617) 864-4810.

Staff: Mike Friedman, Publications Manager; Mary Gilleece, Office Manager; Noe Medina, K12 Consultant; Monty Neill, Managing Director; Sarah Stockwell, Gender/Race Bias Coordinator; Bob Schaeffer, Communications.

Board of Directors: Prof. Denise Carty-Bennia, Northeastern University Law School, Chair; Dr. Pamela George, NC Central University; Paul Pottinger, New Manhattan Corp.; William Robinson, Dean, D.C. School of Law; Bob Schaeffer, Treasurer; Steve Solomon, Esq.; Chuck Stone, *Philadelphia Daily News*; Ron Vera, Esq., Barbosa & Vera; John Weiss, President.

Partial funding for FairTest's work comes from the Ford Foundation and the Rockefeller Family Fund.

Fair Testing Code:
Good Principles, No Teeth

A four-page "Code of Fair Testing Practices in Education" has recently been published by a coalition of professional groups representing testing industry interests. The new Code lists what it calls "the major obligations" of developers and users of educational exams in the areas of "developing/selecting appropriate tests," "interpreting scores," "striving for fairness," and "informing test-takers." Its authors claim the Code is a more popular version of portions of the "Standards for Educational and Psychological Testing" adopted by many of the same groups in 1985.

Because they are general statements, test publishers are free to interpret their meaning. The Code does not mandate that such vital information as content, construct, or criterion-related validity studies be made public—indeed, it does not require that such studies even be done.

Similarly, the Code calls for tests "that are as fair as possible for test takers of different races, gender, ethnic backgrounds or handicapping conditions." But there are few details describing what steps exam developers should take to implement this high-minded goal. Nor are there any provisions specifying what rights test-takers should have through Truth-in-Testing or similar consumer safeguards.

The biggest gap of all, however, is the lack of any enforcement mechanism. Like other professional codes, it's effectiveness depends solely on self-regulation. Unfortunately, the history of testing industry abuses makes this prospect unlikely.

The Code of Fair Testing Practices may be a laudable statement of principles, and at least raises the point that testmakers and users ought to be held accountable for their actions. But it's no substitute for formal public oversight and control.

Single copies of the Code may be ordered for free from the National Council on Measurement in Education, 1230 Seventeenth Street, N.W., Washington, D.C. 20036, and will be printed in the December 1988 FairTest Update on Testing Reform.

Grade School Testing

New Groups Oppose Testing of Young Children

The American Federation of Teachers (AFT), the Intercultural Development Research Association (IDRA) and the Alabama Association of Early Childhood Teacher Educators (AAECTE) have recently joined the growing number of organizations strongly opposed to the use of standardized tests on young children. Their new position statements strongly reinforce recent decisions by Mississippi, North Carolina and Arizona to limit or abandon the use of standardized tests for young children.

In a resolution adopted at its annual convention in July, the AFT declared that standardized tests of kindergartners "may be misused" and have resulted "in too much emphasis on academic skills" in kindergarten. The AFT resolution criticized both the reliability and educational soundness of such tests for kindergartners. According to the September 1988 issue of the AFT magazine *American Educator*, pre-kindergarten screening tests "raise serious questions about equity, since they can have the effect of denying developmentally appropriate experiences to children who can most benefit from them, such as at-risk youngsters."

The resolution also states, "[I]t is not at all unusual to see kindergartens where children work individually on worksheets and skillsheets and engage in other forms of drill and practice. Child specialists contend that this is unsound for all children." The resolution concluded that "(tests) cannot be a substitute for educational and moral decisions" and urged that "more attention should be focused on materials that help make school a joyous experience of success, including those that emphasize active learning, individual experiences, cooperative learning and non-age-graded schooling."

At the summer meeting of its Board of Directors, the San Antonio, Texas based Intercultural Development Research Association (IDRA), an organization focusing on issues related to the education and culture of Hispanics in the Southwest, called for the abandonment of testing in grades K-2. IDRA expressed "severe reservations" about the increase in testing in all grades over the last several years, noting problems with administration, reliability and validity, and racial/ethnic, cultural, class and language bias which exist in standardized tests.

IDRA's statement also questioned the validity of standardized tests for any children below grade three, saying that "achievement test scores at early ages are commonly more indicative of differences in individual developmental patterns of children than of differences in learning efficiency or capacity."

In a March 25, 1988 position paper, the Alabama Association of Early Childhood Teacher Educators (AAECTE) called for a complete moratorium on the use of achievement tests in grades K-3, citing pedagogical problems, teaching to the test and pressures on teachers to produce higher test scores.

Severely critical of the underlying assumptions of achievement testing, that reading and writing are "isolated bits of knowledge children acquire by association" and that "mathematics is...a collection of symbols and conventions," the AAECTE position paper calls on educators to "reconceptualize instructional objectives," and to use appropriate alternative developmental assessments instead of standardized, multiple-choice tests.

The AAECTE has asked the Alabama Board of Education, State Department of Education and local school superintendents to declare a moratorium on standardized testing in grades K-3.

Testing in Boston Blasted

The Way Out: Test Score Patterns and Reading in Boston Public Schools, a Status Report from the Massachusetts Advocacy Center (MAC), condemns the use of inappropriate reading tests and release of misleading test score results by the Boston Public Schools. The report concludes that the city's overreliance on testing is educationally unsound.

To support its charges, MAC cites remedial reading programs which "are little more than test preparation classes with the resulting rise in test scores achieved at significant cost, both in financial terms and in terms of time robbed from real reading instruction." MAC called for significant changes in the structure and processes of education, and concluded: "As long as Boston continues to equate higher test scores with real school improvement, needed reforms will be delayed."

(The report is available from MAC, 76 Summer St., Boston, MA 02210; $2.)

Vermont

Vermont is considering the use of alternatives to standardized tests as part of proposed new statewide evaluations. See the December 1988 *FairTest Update on K-12 Testing* for details.

Grade School Testing

Test Coaching for Babies

The Baby Boards: A Parents' Guide to Preschool and Primary School Entrance Tests, by Jacqueline Robinson (New York: Arco (Prentice Hall), 1988, 223 pp, $12.95)

If you want to help your child train for one of the standardized entrance exams that are becoming increasingly common for pre-school or kindergarten admission, this might be the book for you. But if you want to challenge the use of these exams, it won't be of much help. The message of *The Baby Boards* is that at least some of these tests are or can be in children's best interest.

Part I discusses in some detail two of the most common "baby boards," the Wechsler and Gesell tests. Ms. Robinson presents basic information about what is on the exams, how they are administered and scored, and questions parents should ask. Part II explains how to prepare your child for each test and includes many exercises to help you teach your child the content the tests cover.

Ms. Robinson does touch on some dangers associated with standardized tests. She points out that a child deemed "not ready" for kindergarten or first grade could spend an extra year in school. She also notes that in California "all IQ tests have been banned as a means of diagnosing or placing black children," and observes that IQ tests have been "abused more than any other [tests]." Moreover, her tone is clearly skeptical regarding the claims made by Gesell of a standard developmental progression that their tests can measure. (For more information on the Gesell, see story on page 5.)

But these cautions are merely noted in passing in what is otherwise a test prep book that accepts the controversial assumption that testing young children is reasonable. The underlying psychological, developmental, social, cultural, racial, class and gender assumptions embodied in the tests are never questioned.

Even a cursory look at the reviews of these products in the *Mental Measurements Yearbooks* and other professional publications indicates the small norming samples, the absence of validity studies in most significant areas, and the limited and often outdated psychological and developmental assumptions inherent in the tests. *The Baby Boards* ignores the fact that the results of many of these instruments are inaccurate and misleading for young children, par-ticularly those who are not white, middle to upper class. Nowhere does the book mention the growing number of organizations that oppose the testing of young children.

Ms. Robinson might acknowledge many of these criticisms, but reply that because the tests are given so often and the fate of so many children tied to them, it is important that parents know how to prepare their children.

Such an argument, however, does not excuse the form of presentation. All the book's graphics are of White people who are clearly of a middle to upper class background, and these graphics reflect the style and content of the text. Instruction in exam routines, no matter how absurd or downright wrong the test exercise may be, is presented as fun games to be played with the child. Test administrators are presented as warm, friendly and not hurried, the parents as informed and helpful, the school as an inevitably delightful place. This clearly is not the situation for too many children in too many schools.

It would seem that the author was aiming for a limited audience: the people whose children are already most likely to do well on the test. Though financially comfortable White parents have the most power to demand that schools stop the misuse of tests, the book does not suggest that parents ought to fight against "baby boards" individually or in groups. Given the profound flaws in these tests, battles to stop their use could be as successful as the campaign to stop standardized testing in first and second grades has been in North Carolina.

In sum, if you want a set of "games" with which to prep your child and some basic information on how the tests are structured, this book can be useful. If you want even minimal critical information so you can help prevent the abuse of your child and other children by tests and testers, *The Baby Boards* is not helpful. That's too bad, because a good book would have done both.

Nebraska

Education Week of September 28, 1988 reported that the Nebraska Board of Education has approved a new method of accrediting public schools, which includes evaluating students' college admissions test scores. In addition to ACT and SAT scores, the new formula will examine dropout rates, student performance after graduation and may also include a survey of students and parents to see if they are satisfied with schools.

The new guidelines also call for new state mandated testing of a sample of students in order to "gauge progress." These new rules replace the old method of accrediting schools which included such things as student-staff ratios, size of the school library and "whether teachers are properly prepared for the classes they teach." The new rules go into effect July 1, 1989.

Grade School Testing

Challenge to Gesell as Kindergarten Placement Exam

Use of the Gesell School Readiness Test to place children in kindergarten has been successfully challenged by an upstate New York family.

The parents in the case were taken by suprise by a notice from the Norwood-Norfolk School District that their daughter would have to undergo testing for placement in kindergarten. Using the fifteen minute long Gesell test, which claims to determine a child's "developmental age," the plaintiff's daughter was deemed "unready" for regular kindergarten, and was placed in a "developmental" kindergarten.

Upon investigation, however, the parents discovered that 61% of the children in the Norwood-Norfolk district were found to be unready for regular kindergarten. Most of the children who finished the developmental kindergarten were then placed either in regular kindergarten or a so-called "pre first grade" program. During the 1986-87 school year, only two children were moved from the developmental kindergarten directly into a regular first grade class. This led to many children not entering second grade until age 8, a year behind their peers.

After meetings with school and district officials resulted in no change to their daughter's status, the parents went to a legal services attorney, Linda Mills, who filed a class-action brief with the State Department of Education on behalf of all the affected students. The complainants alleged that use of the Gesell to place students in developmental kindergartens violated several state and federal laws, including the New York statute giving children the right to enter kindergarten at age 5.

Mills claimed that since the school district was treating the affected children as if they were handicapped, they were entitled to the protections afforded by laws regulating the education of handicapped children, including section 504 of the Rehabilitation Act of 1973, the Education for All Handicapped Children Act and the New York Special Education law. These statutes guarantee individual evaluations and educational programs, procedural due process and possibilities for placement in regular kindergarten classes.

In affidavits accompanying the parents' complaint filed with the State Department of Education, researchers from the University of Michigan, the University of Colorado and the State University of New York (SUNY) at Albany all were heavily critical of the Gesell.

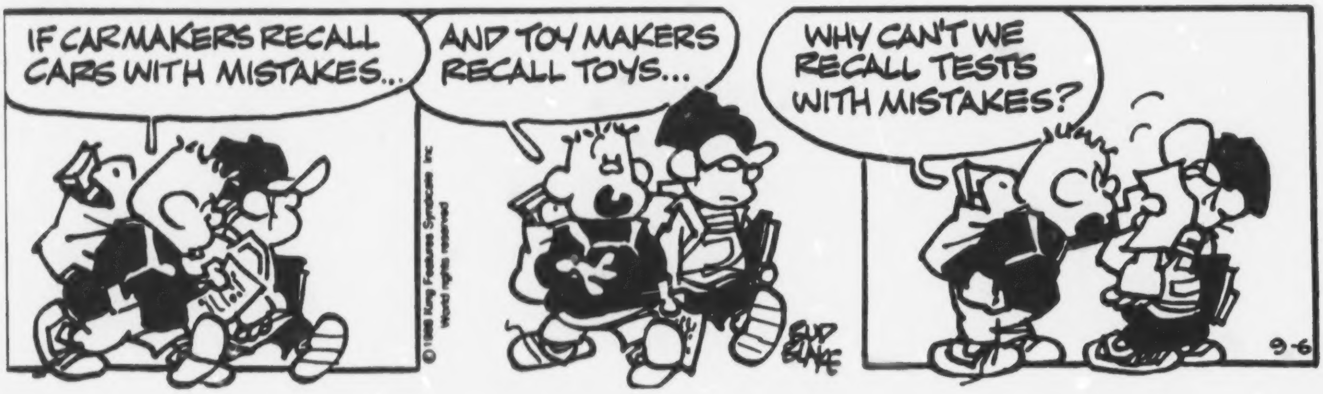
Dr. Samuel Meisels of Michigan wrote: "...Other information (besides test scores) is required before a valid judgement can be made about whether a child should...effectively be denied a free and appropriate public education on the basis of unsound test labels..." Dr. Lorrie Shepard of Colorado said that the extremely high percentage of children placed in developmental kindergarten "calls into question the validity and reliability of the test."

Dr. Deborah C. May of SUNY-Albany reported the results of her study of children placed in developmental kindergartens that found that these children were more likely to score below their peers in later years than those who had been in regular kindergarten. May's study also showed that children whose parents had removed them from developmental kindergartens were more successful in school than those who had been through a developmental program.

According to reviews contained in several recent editions of the prestigious *Mental Measurements Yearbook*, the Gesell School Readiness Test was normed on a sample of only 640 children aged 2-6

(continued on page 6)

IGER by Bud Blake


IF CARMAKERS RECALL CARS WITH MISTAKES...
AND TOY MAKERS RECALL TOYS...
WHY CAN'T WE RECALL TESTS WITH MISTAKES?
© 1988 King Features Syndicate Inc. World rights reserved
BUD BLAKE
9-6

Grade School Testing

Alternative Evaluations Under Chapter I?

Federal standards for funding remedial math and reading programs under Chapter I have been overhauled by the School Improvement Amendments of 1988. Parents now must be involved in determining local goals and program structure, including evaluation procedures.

The House and Senate conference committee report recognized "the congressional intent that measures and standards used to demonstrate progress toward desired outcomes may be something other than nationally normed standardized test scores."

However, the Department of Education's draft regulations covering "National Evaluation Standards" mandate the use of standardized tests to measure student achievement. This could force programs to teach to the test, thereby undermining local program goals.

The proposed regulations do state that schools also may use other "indicators of program effectiveness" to measure progress toward local goals. *Examiner* readers should respond to the draft regulations by *December 20* to urge that the regulations more strongly encourage and value alternative academic assessment. (The draft regulations are in the *Federal Register*, Vol. 53, No. 204, Oct. 21, 1988, pp. 41466-41492.)

The creation of viable alternatives for Chapter I could have an important effect on the eventual replacement of standardized tests by useful, valid and unbiased instruments. But parents will have to fight to make alternatives more important than standardized tests in evaluating the progress of both children and programs.

Gesell *(from page 5)*

years, nearly all of whom were white and lived in Connecticut. In addition, the *Yearbook* noted that the makers of the Gesell did not provide validity studies to demonstrate that the test measures what it says it measures, and that test administrators may have as little as one day's training.

After the plaintiffs refused a settlement that would have affected only their child, the school district gave in and reached a settlement covering all the children who had been placed in developmental kindergartens. The arrangement provides for drastic changes in the kindergarten placement system in the Norwood-Norfolk school district, including the parents' right to place their child in regular kindergarten, provision of test names and scores used in placement decisions, and parental notification of the right to refuse placement of their children in developmental kindergarten.

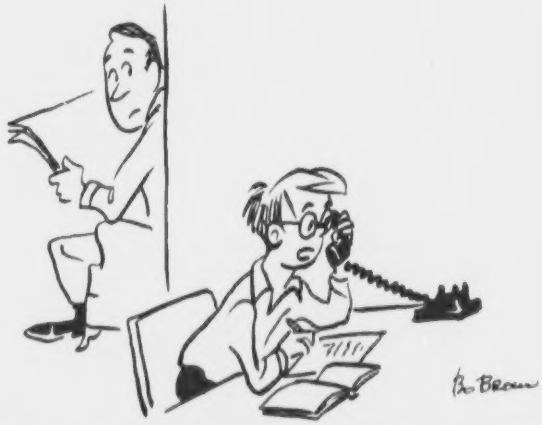
(For copies of the pleadings in this case please contact the Harvard Center for Law and Education, Larsen Hall, 14 Appian Way, Cambridge, Mass. 02138; (617) 495-4666.)

California Cheating *(from page 1)*

that they were teaching to the test and placing too high an emphasis on increasing test scores.

"We are asked why the kids are not performing to grade level... Lately we've had to devote 20-30 minutes a day to prepare the students," Ilene Weise of Los Angeles told the *Los Angeles Times*. Wayne Johnson, president of United Teachers-Los Angeles said, "Teachers are told if you don't raise your scores, you are a bad teacher." But others felt less pressured: "In my classroom, testing is no big deal," said Julie Cornell of Los Angeles. "I don't personally buy into the theory that testing helps the students or builds self-esteem," she added.

In the past several years, cheating scandals have been reported in Virginia, Ohio, Maryland and New York. All stem from the belief that standardized tests are the panacea for complex educational problems. As long as schools and states continue to put heavy emphasis on test scores, scandals like these will continue to erupt periodically.



"When I ask you to give me the answers to the math homework, Albert, just look on it as a sort of insider trading."

NTE Questions and Answers

(from page 16)

Science

30. The very large number of life forms on Earth has led biologists to develop a system of classification that brings some order out of the diversity. This ordered system is most similar to which of the following?

a) The English alphabet
b) The Roman alphabet
c) A federal income tax form
d) An address on an envelope
e) A Romance language

Comment: Correct answer (d) was selected by 20% of the examinees. For this science question, the "basic skill" is precise knowledge of the structure of a "proper" business envelope address, hardly necessary for one to be a good science teacher . In fact, this too is a bad answer unless they use an inverted definition of "similar." Biological classification goes from the general (phylum) to the specific (sub-species), while a "proper" envelope goes from specific (recipient's name) to general (city/state/zip) or even (country) in the case of foreign mail. What to the NTE is "similar" is in fact opposite.

Defective Questions

These questions are not isolated examples. An internal review by ETS has agreed with critics who claim that many questions on many parts of the NTE Professional Knowledge and Communication Skills, as well as General Knowledge, either have no "right" answer or include more than one correct answer choice. ETS recently admitted that 11 questions on exams given between 1983 and 1987 contained such flaws.

ETS spokespeople bragged that the errors totalled less than 1% of all items. However, critics have found a much higher error rate. Dr. Lorrie Shepard of the University of Colorado says close to 40% of the questions on one form of the Test of Professional Knowledge had either no correct answer or more than one correct answer.

Even with ETS's limited rescoring of the 11 admittedly flawed questions, the scores of 1,200 test-takers were raised. At least 19 prospective teachers who had previously failed to earn passing marks now meet state licensing requirements. But the ETS review and the broader discussions about the number of erroneous "correct" responses do not address a larger issue about the test: *there is still no evidence that the NTE can accurately predict who will and will not become a competent teacher.*

Stop NTE Use Now

The NTE never was an unbiased, basic skills test accurately measuring the core knowledge that all teachers—English, math, science, remedial, kindergarten and shop—had to know. Instead it is an exam similar to the SAT that draws on irrelevant and often culturally-specific facts to spread examinees across a wide range of scores.

By claiming that the NTE measures basic skills, ETS implicitly argues that someone who does not pass the test cannot be a good teacher. This claim has never been proven in a study demonstrating predictive validity. Given the exam's bizarre content, it's hard to see how failing the test precludes the possibility of someone becoming an effective teacher. The question is now obvious: *Why must four more classes of prospective teachers take an exam even its manufacturer admits is inadequate?* How does ETS explain its failure to withdraw a defective product even sooner?

A Better NTE?

The new proposed verson of the NTE will have a "general knowledge" test given as early as the sophomore year of college (an equivalent to the PPST?). Will it be more relevant, more carefully constructed, less obscure and culturally biased, than the current NTE? Will it be, in fact, a "basic skills" test? If so, will ETS do some real predictive validity studies?

The new exam will also include a "professional knowledge" test. Will it have fewer errors than the 40% that Dr. Shepard found? Will the new "professional knowledge" as well as the new "subject matter" and "pedagogy" tests be free of the cultural, racial, class and gender biases that have plagued past NTE exams?

In sum, what safeguards will ETS impose to insure that the replacement for the NTE is fair, valid and relevant? The track record of ETS in defending the biased and flawed NTE indicates that the new test will require the strictest outside scrutiny.

US Agency Finds Texas Test Unfair

The Federal Equal Employment Opportunity Commission (EEOC) has ruled that the 1986 Texas "literacy test" discriminated against teachers who were Black or over age 40. The state required all classroom educators to pass the test in order to keep their jobs.

The EEOC concluded that the Texas Examination of Current Administrators and Teachers (TECAT) "adversely impacts on blacks and persons within the protected age group." No remedy has yet been ordered.

State officials say they disagree with EEOC and are contemplating a challenge. Lawyers for some of the 1,875 teachers who failed the test are also considering legal action to enforce EEOC's judgement by seeking reinstatement and back pay.

University Admissions Testing

NMSC *(from page 1)*

test scores not be high when they are the *only* factor used to determine initial eligibility for the awards?

- NMSC and ETS imply that the signers of the public letter advocate "quotas," another favorite straw man. In fact, the public letter does not support quotas. But the NMSC itself already manipulates test results to affect the number of scholarship winners. NMSC uses a "Selection Index" (doubling the PSAT-Verbal score and then adding the PSAT-Math score) to increase the number of young women in the semifinalist pool and uses different qualifying scores for each state to assure geographic diversity. This clearly demonstrates that even NMSC recognizes that the PSAT/NMSQT is inadequate as the sole factor in determining semifinalists.

- NMSC and ETS continue to ignore the growing body of evidence proving bias in the SAT, the test from which the PSAT is derived. Despite FairTest's report *Sex-Bias in College Admissions Tests: Why Women Lose Out*, numerous articles in the *FairTest Examiner*, books like *The Case Against the SAT*, and even ETS internal research on bias against women and minorities, the NMSC/ETS memo claims that "FairTest does not provide any evidence to support" the contention that the PSAT/NMSQT is a flawed test. (A "Selected Bibliography on Unfair Bias in the SAT" is now available free of charge from FairTest; send request and SASE.)

NMSC's strategy is not working. Organizations, corporations and even NMSC board members are increasingly expressing doubts about NMSC's current formula. In a letter to NMSC, Andres Tobar, Education Committee Chair of the League of United Latin American Citizens declared, "LULAC does not condone the use of [the PSAT/NMSQT] because it discriminates against Hispanics in what the test measures, in the manner it is administered and in its results.... We are not requesting that awards be given to Hispanic students solely on the basis of the numbers they represent in the schools. We are asking that the methods used to measure the academic capabilities of America's students be such that they allow for Hispanics to compete fairly."

Fred Azcarate, President of the United States Student Association (USSA), also criticized NMSC's use of the test, pointing out in a letter to NMSC: "Schools and corporations generously provide scholarships to individuals based on NMSC's assessments. America's students and the NMSC scholarship sponsors deserve a program which determines recipients fairly."

Corporate funders of National Merit Scholarships have raised similar issues. One wrote to FairTest, "I share your concerns... regarding the merits of the PSAT/NMSQT," while another asserted, in a letter to NMSC, that "(t)he program should provide students the same access to scholarships regardless of race or sex. The allegations of unfairness expressed by FairTest should be evaluated."

The National Merit Scholarship Corporation would be wise to take action before the chorus grows even larger and louder.

NMSC Signer List

In less than four months nearly 30 more groups and three additional members of Congress have joined the original 18 signers of FairTest's public letter to the National Merit Scholarship Corporation (NMSC). Organizations as diverse as the Lawyer's Committee for Civil Rights Under Law, the National Association of Independent Schools (NAIS), the National Congress of American Indians, the Women's Law Project, the Puerto Rican Legal Defense and Education Fund (PRLDEF), the National Institute for Women of Color, and the Japanese American Citizens League have joined the effort to convince NMSC to change its formula.

The current signers include:

FairTest, The National Center for Fair & Open Testing
A Better Chance, Boston
Asian American Legal Defense and Education Fund (AALDEF)
Asian and Pacific Americans Civil Rights Alliance
Asian Law Alliance
Asian Pacific Law Center
ASPIRA Association
Center for Women Policy Studies
Equality in Testing Project, Phyllis Rosser, Director
Japanese American Citizens League
Japanese American Citizens League, New York Chapter
Korean-Americans for Peace and Justice
Labor Council for Latin American Advancement
Lawyer's Committee for Civil Rights Under Law
League of United Latin American Citizens (LULAC)
Mexican American Legal Defense and Education Fund (MALDEF)
Mexican American Women's National Association (MANA)
National Association for Bilingual Education (NABE)
National Association for Ethnic Studies

University Admissions Testing

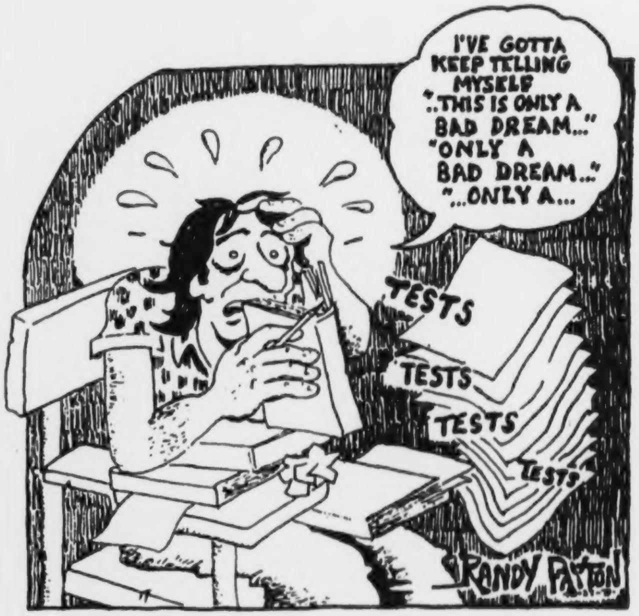
NACAC Urges Proper Use of Test Scores

The National Association of College Admissions Counselors (NACAC), which has about 1100 institutional members in higher education, recently approved several additions to their "Statement of Good Principles and Practices" condemning misuse of standardized tests. The Principles now call on NACAC members to evaluate and justify their use of test scores in the admissions process and to avoid using test scores as the sole criterion for admission to college. The guidelines on test use were the only substantive changes to the "Principles," which had not been revised in several years.

The admissions counselors further declared that test scores should not be used in making financial aid decisions, particularly for students whose test scores may reflect deficiencies caused by their socio-economic status, race, gender or ethnicity. Also, published profiles of institutions should include statistics on first year students who actually enrolled, rather than all students who applied or all students admitted.

When an institution is thought to be in violation of the "Statement of Good Principles and Practices," a complaint is filed, and a committee judges the complaint. According to NACAC Executive Director Frank Burtnett, when a college is informed that they are in violation of NACAC guidelines they generally change their practices to eliminate the violation. The worst penalty that violators can be slapped with is expulsion from the organization, and the notoriety that goes with that expulsion.

In addition to test use, the "Principles" cover such areas as marketing, how schools present themselves, financial aid, application deadlines and conduct at college fairs.


I'VE GOTTA KEEP TELLING MYSELF "..THIS IS ONLY A BAD DREAM..." "ONLY A BAD DREAM..." "...ONLY A...
TESTS
TESTS
TESTS
TESTS
RANDY PATTON

More Than Doubles

National Association for the Advancement of Colored People (NAACP)
National Association of Asian American Professionals
National Association of Independent Schools (NAIS)
National Black Women's Political Leadership Caucus
National Conference of Puerto Rican Women
National Congress of American Indians
National Indian Youth Council
National Institute for Women of Color
National Organization for Women (NOW)
NOW Legal Defense and Education Fund/Project on Equal Education Rights (PEER)
National Political Congress of Black Women, Shirley Chisholm, National Chair
National Puerto Rican Coalition
National Society of Black Engineers
National Women's Law Center
National Women's Studies Association
New York Public Interest Research Group (NYPIRG)
ORBIS Associates/Indian Education Resource Center One
Puerto Rican Legal Defense and Education Fund (PRLDEF)
Sage Women's Educational Press
Eleanor Smeal
Sociologists for Women in Society
U.S. Delegate Walter E. Fauntroy (D-DC)
U.S. Representative Daniel K. Akaka (D-HI)
U.S. Representative Norman Y. Mineta (D-CA)
U.S. Representative Major R. Owens (D-NY)
U.S. Representative Claudine Schneider (R-RI)
U.S. Students Association (USSA)
Women's Equity Action League (WEAL)
Women's Law Project
Women's Legal Defense Fund

University Admissions Testing

New York Scholarship Reversal

Due to legislative red tape, New York State has temporarily dropped its one year old reform that had greatly increased the percentage of women chosen to receive $40 million in state funded scholarships.

Responding to charges that the old system, based solely on SAT or ACT scores, gave an unfair advantage to boys, the state legislature had adopted a 1/2 grades, 1/2 test score formula for use in the 1987-88 school year. The reform's impact was to be reviewed after one year. Under the new system, using high school grades in combination with test scores, girls won 50% of the basic Regent's scholarships and 40% of the top Empire awards, large increases from the previous figures of 43% and 28% respectively.

But many educators complained that the state Department of Education sabotaged the new scholarship process by sending out contradictory and confusing memos to high schools explaining the new system. As a result, grades were not reported consistently.

Nevertheless, the state legislature was poised to endorse the system by extending the "1/2 grades, 1/2 SAT scores" formula for another year when it recessed for the summer because of conflicts over unrelated issues. Therefore the "100% SAT score" formula was reinstated by default.

But even the test-maker (ETS) and sponsor (College Board) are opposed to this system. Fred Moreno, a spokesperson for the Board, said "The College Board does not support the use of test scores as the sole basis for important decisions affecting the lives of individuals."

In a recent interview with the *Albany Times-Union*, ETS President Gregory Anrig told reporter Tim Spofford that New York should revive the new scholarship formula. "The two (grades and SAT scores) together are better," Anrig said.

Such high-sounding statements are laudable, but they don't go far enough. FairTest and the New York Public Interest Research Group (NYPIRG), the organization which led the coalition that won the 1987 reform, have called on ETS and the College Board to withhold scores from the New York Scholarship Agency so long as it uses a system which violates guidelines for proper test use. The test companies' responses will prove the sincerity of their rhetoric.

NCAA Prop. 48 Discriminates Against Blacks

Proposition 48, the National Collegiate Athletic Association (NCAA) regulations governing college admissions for student athletes, has a hugely discriminatory impact on black student athletes. The regulations require student athletes to maintain a 2.0 grade point average in a high school core curriculum and score at least 700 on the SAT or 15 on the ACT. A recent survey by the Associated Press

(continued on page 11)

Massachusetts Scholarships Using SAT Improperly

After uncovering the misuse of test scores in awarding New York State and National Merit scholarships, FairTest found another example of blatant test misuse right under its nose: the Massachusetts Honors Scholarships are given out solely on the basis of SAT scores. The all-too-typical result: More than 60% of awards went to males in 1988. Though ethnic breakdowns are not available, the racial bias in the test ensures that the number of Black and Hispanic winners must be minimal.

The scholarships, which pay for one year's tuition at any public four-year school in Massachusetts, are given out to the top four SAT scorers who apply from each state senatorial district. In addition to relying on the SAT as the sole factor in making "high stakes decisions," a misuse even ETS officially opposes, a number of other bureaucratic details make this system even more unfair. Students who don't take the SAT, overlook the scholarships listed on the registration materials, or can't use one of their four free score reports to send scores to the Massachusetts Scholarship Commission, aren't eligible for the awards.

To make matters worse, the program applies a "floor" below which a winner's score cannot fall. So even if a student is the valedictorian of her class, has a straight "A" average and participates in many outside activities, she is not good enough for a Massachusetts Honors Scholarship unless her score is above the cut-off.

FairTest has written a letter to ETS President Gregory Anrig demanding that ETS withhold SAT scores from the Massachusetts Scholarship Commission because of this blatant abuse of scores.

University Admissions Testing

Proposition 48

(continued from page 10)

revealed that, similar to past years, 87% of the freshman football players disqualified from playing in 1988 were Black.

As a result, student athletes, many of whom depend upon their scholarships to finance their educations, may be denied access to college because of racially discriminatory tests.

But worse is yet to come: currently, if a student athlete fails to meet one of the requirements, he or she can still be admitted to college on an interim basis, but is ineligible to play during the first year. In the future, athletes who fail to meet both Prop. 48 standards may be denied entrance to schools in the Southeastern Athletic Conference (SEC), which includes such schools as the Universities of Alabama and Georgia.

At their summer meeting, the presidents of the SEC schools voted to introduce legislation at the January 1989 NCAA convention to toughen college eligibility standards under Proposition 48. The SEC also decided that by 1993 member schools will no longer admit athletes who fail to meet both the test score and grade point average standards.

Also this summer, the Colonial Athletic Association decided to submit a proposal to the NCAA convention that would deny athletic scholarships to athletes who fail to meet one of the requirements of Prop. 48. Such provisions would likely have the effect of denying a college education to many minority students.

Yet the test score portion of Prop. 48 is invalid for predicting who can succeed in college. A study conducted by the *Ann Arbor News* (Michigan) found that 79% of football players who were denied freshman eligibility at the largest football playing universities in the country were still enrolled at their institutions.

This is on top of clear evidence that the SAT underpredicts minority athletes' abilities. Professor Timothy Walter of the University of Michigan, who conducted a study of student athletes at Michigan said: "SAT scores consistently predict failure when in fact kids succeed." (See *FairTest Examiner*, Vol. 2., No. 1.)

Since the tests do not predict success and discriminate against Blacks, they should be removed as barriers to freshman eligibility, not made barriers to enrollment and financial aid by the NCAA.

THURSDAY, SEPTEMBER 22, 1988 · USA TODAY

This test is an unfair measure

By Sarah Stockwell
Guest columnist

CAMBRIDGE, Mass. — It makes no sense to measure educational health with a broken thermometer. Yet that's what the College Board is trying to do by focusing on tiny year-to-year changes in Scholastic Aptitude Test scores. Its public-relations campaign ignores the real issues about the use of the SAT.

Who cares about one- or two-point differences when the test is fundamentally flawed? Scores of studies show the SAT is biased against women, minorities and students from low-income families. Questions regularly deal with such upper-class activities as regattas, polo, minuets and horseback riding. None measure skills like creativity, writing abilities or motivation. As a result, SAT scores are nearly useless at measuring college aptitude or high school performance.

Women score much lower, even though they earn higher grades than men in both high school and college. Many minorities succeed in college despite the SAT's prediction of failure.

Coaching compounds the flaws. More than a dozen studies prove prep courses can increase scores 100 points or more. Families who can afford $500 to $600 for coaching can give their children another leg up. But a young black woman from a low-income family faces triple discrimination: race, gender and parents' inability to pay for coaching.

Sarah Stockwell is admissions test coordinator, FairTest: National Center for Fair & Open Testing.

In part because of these biases, the SAT has essentially no value in admissions. The New York Public Interest Research Group found high school grades are three times better at predicting college performance. Yet forecasting freshman grades is just what the test is supposed to do, according to its manufacturers.

The real issue is why 1.5 million students a year are forced to take such a defective test. Why should families spend $100 million annually on fees and test preparation? Why should National Merit Scholarships use an SAT clone as qualifying hurdle for $40 million in awards? These are the questions the College Board wants ignored.

The truth is that the SAT simply doesn't work — unless you are female, non-white or poor, and then it works against you. Unless the test is thoroughly overhauled, the SAT should be discarded.

Gifford Commission

In September, "fifteen educational, political and corporate leaders" were named as members of the National Commission on Testing and Public Policy. The Ford Foundation-funded Commission is charged with making a comprehensive review of the role of testing in America.

Commission chair Dean Bernard R. Gifford of the Graduate School of Education of the University of California at Berkeley has commissioned a number of papers and held several hearings at which representatives from FairTest and other groups testified. The *Examiner* will report on the Commission's work as its studies are released to the public.

The members are:

- Jose A. Cardenas, executive director, Intercultural Development Research Association, San Antonio, Texas
- Julius L. Chambers, director-counsel, NAACP Legal Defense Fund, New York
- Frederick R. Chang, director, Information and Decision Sciences, Pacific Bell, San Ramon, California
- Gov. Bill Clinton (D-Arkansas)
- Douglas A. Fraser, former president, United Auto Workers, now University Professor of Labor Studies, Wayne State University, Detroit
- Raymond C. Fisher, past president and member, Los Angeles Civil Service Commission
- Badi G. Foster, president, Aetna Institute for Corporate Education, Hartford, Connecticut
- George H. Hanford, president emeritus, College Board, New York
- Katharine H. Hanson, executive director, Consortium on Financing Higher Education, Cambridge, Massachusetts
- W.W. Herenton, superintendent, Memphis City Schools, Memphis, Tennessee
- Antonia Hernandez, president/general counsel, Mexican American Legal Defense and Education Fund, Los Angeles
- Elridge W. McMillan, president, Southern Education Foundation, Atlanta
- Patsy Takemoto Mink, former member, U.S. House of Representatives, Honolulu
- Edward E. Potter, attorney, McGuiness & Williams, Equal Employment Advisory Council, Washington, D.C.
- Thomas G. Sticht, president, Applied Behavioral and Cognitive Sciences, Inc., San Diego.

Quote of the Month

"This is not a golf game. We can't reduce performance to a single numerical score."

—*Richard P. Mills,*
Vermont Commissioner of Education

Tests Driving You and Your Kids Crazy?

Order your copy of FairTest's best-selling report *Fallout From the Testing Explosion: How 100 Million Standardized Exams Undermine Equity and Excellence in America's Public Schools*, now in an expanded second edition.

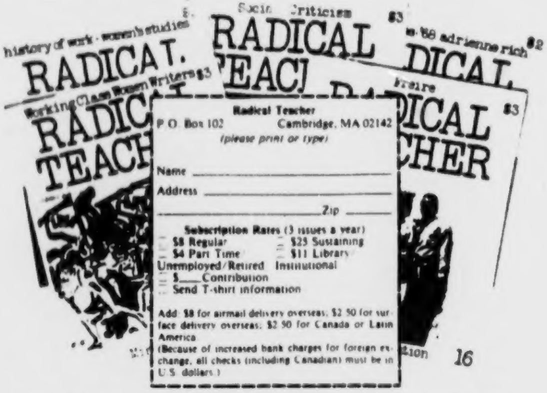
The report includes the results of FairTest's survey of every state and selected school districts around the country. The expanded edition includes new data on validity of tests, and severe critiques of several of the major test batteries.

To get your copy of *Fallout From the Testing Explosion*, send a check for $8.95 (or an institutional purchase order) to:

FairTest, P.O. Box 1272,
Cambridge, MA 02238-1272.

RADICAL TEACHER

- "**Radical Teacher's** activist perspective confronts the real issues of classroom and community . . . A practical and valuable journal, based on solid research and theory." *The Guardian*
- Previous issues on: Back-to-Basics, Lesbian and Gay Studies, Health Sciences Education, The Politics of Literacy, Mass Culture, Black Women's Studies, Marxist Teaching, Feminist Pedagogy
- **Radical Teacher** is an independent socialist and feminist magazine that focuses on three critical areas: the politics of teaching; the political economy of education; feminist, Marxist and third world perspectives in literature, history, biology, sociology and other disciplines



Radical Teacher
P.O. Box 102 Cambridge, MA 02142
(please print or type)

Name ________________
Address ________________
________________ Zip ________

Subscription Rates (3 issues a year)
☐ $8 Regular ☐ $25 Sustaining
☐ $4 Part Time ☐ $11 Library/
Unemployed/Retired Institutional
☐ $____ Contribution
☐ Send T-shirt information

Add: $8 for airmail delivery overseas; $2.50 for surface delivery overseas; $2.50 for Canada or Latin America
(Because of increased bank charges for foreign exchange, all checks (including Canadian) must be in U.S. dollars.)

Add: $8 for airmail delivery overseas; $2.50 for surface delivery overseas; $2.50 for Canada or Latin America
(Because of increased bank charges for foreign exchange, all checks (including Canadian) must be in U.S. dollars.)

Internal Visiting Committee Criticizes ETS Test Misuse

As part of its continuing public relations effort for the SAT and its other tests, the Educational Testing Service appoints a committee every year to study the company and make recommendations for changes in policy and practice that will make sure that ETS adheres to its "Standards of Quality and Fairness." But sometimes ETS gets more feedback from the Visiting Committee than it planned for. The following is a verbatim quote (emphasis added) from p.8 of the 1987 Committee report:

"...the Visiting Committee has deep concerns. There continue to be genuine instances of the misuse of tests and a national climate that almost guarantees that this will be an increasing problem in the future. At the same time, *there is little sense of informed corporate awareness at ETS with respect to a serious and continuing misuse of its tests and those of its clients*...ETS needs to take the initiative in providing leadership on problems associated with broad and systematic test misuse.

"It is the conviction of the Visiting Committee that ETS has a public responsibility to assume a more proactive posture in identifying and acting on misuse of its own tests and in raising professional, public and constituency consciousness on these issues. Specifically, ETS should move aggressively to carry out at the corporate level the principles and policies contained in the Test Use and Public Information sections of the revised 'ETS Standards for Quality and Fairness.'

"As things presently stand, *there is inadequate accountability at ETS for problems of broad and systematic test misuse*. The Visiting Committee believes it is necessary to create an independent entity within ETS with specific responsibility to coordinate activities in this area and to ensure that ETS speaks effectively on the issue."

In addition to criticizing general test misuse at ETS, the company's National Teacher Exam (NTE) programs also came under heavy fire from the 1987 Visiting Committee. The Committee repeated several questions which have been raised by FairTest and testing reform activists over the past several years:

"• Are the NTE Programs contributing to the improvement of teacher education and practice?

"• Do the NTE Programs have a negative effect on the recruitment of minorities into the teaching profession?

"• Are the NTE Programs tests valid for the inferences commonly drawn from their uses?

"• Can ETS continue to rely on studies undertaken by state level clients using the NTE Programs as their primary source of validity?"

Unfortunately, the Committee leaves all of its questions tantalizingly unanswered. By sticking to such generalities as "...the Committee's concern is a recognition that there is a growing pressure for new methods of teacher evaluation," and suggestions "...that ETS take leadership now in developing new approaches to teacher certification and licensure," they fail to address the important concerns of testing reform and civil rights activists about teacher testing.

The Committee recognizes the efforts ETS is making toward creating a replacement for the NTE, but fails to push the company to answer the legitimate questions about validity and misuse of the NTE and the wholesale elimination of minority teachers because of biased teacher licensing tests.

The recommendations the Committee makes for new directions in teacher testing include calls for "new techniques and new approaches" to licensing teachers. Unfortunately, this does little for the thousands of teachers who will be faced with the NTE between now and 1992.

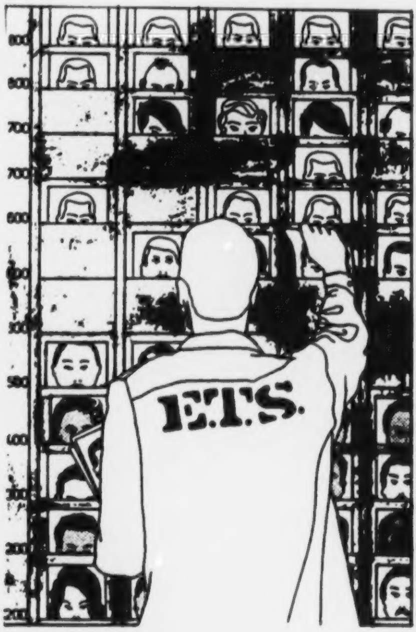


Illustration: Teddy Sandoval ©1979 Equal Opportunity Forum Used with permission.

(The Report of the 1987 ETS Visiting Committee is available free from ETS Publications, 666 Rosedale Road, Princeton, NJ 08541.)

Elementary Principals Issue Testing Recommendations

The Massachusetts Elementary School Principal's Association has issued *Recommendations for the Use of Standardized Tests*. Their proposals include strong statements against test driven curriculum and for limiting the testing of students to once a year, relying on recommendations and teacher-made tests to make important decisions about students, and opening up test materials to public scrutiny.

(The *Recommendations* can be obtained by writing MESPA, 85 Myrick Lane, Harvard, MA 01451.)

Chicago Sun-Times

WEDNESDAY, SEPTEMBER 14, 1988

COMMENTARY

Standardized testing is child abuse

Dennis Byrne

The kids are back in school, which means that a uniquely contemporary form of child abuse is about to begin again—standardized testing.

Parents know about it—the deluge of tests measuring their kids' achievement, intelligence, competence, aptitude, basic skills, vocational interests, remedial needs, predicted college performance, and on and on. At least 100 million are given annually in public schools alone, not including the ubiquitous intelligence tests. While people my age generally took these tests only to get into college, they increasingly are being given to younger children, including kindergartners.

If you're a kid, you know how you compare with every kid in the country, and not in just general terms. You know within a fraction of a percent how many students in your school, community and country can spell better, read better and do square root better.

Such information is of virtually no use to teachers, who should already know what kids need help. It's of no use to parents, except perhaps to increase anxiety levels.

And it damn sure is of no use to kids, who, thanks to the tests, are facing competitive pressures not experienced by most adults. Even before some children master the small motor skills needed to control the pencil used in the test, they risk getting labeled as something they may not be, a label they could spend a lifetime trying to shed. Instead of helping children learn the self-esteem so necessary for later success and satisfaction, the tests merely help children more clearly define their own perceived incompetence.

I guess I'm approaching this from a wrong assumption—that kids should undergo mass testing only if it is somehow good for them. Instead, the tests are good for the unregulated, multimillion-dollar-and-growing testing industry and for education administrators who need something to do. And it is good for self-assured OFs (old fools) who love to engage in generation bashing by citing the latest batch of test scores as evidence of young people's ignorance—not realizing, in their own ignorance, that the tests are unreliable for cross-generational comparisons.

In fact, studies have shown that the tests are not good predictors of performance, and that they are biased against women, minorities and low-income students. Further, they "drive" school curricula, encouraging classroom concentration on tested basic skills at the expense of untestable, higher skills needed later—things like world understanding, applying background experience to new situations, and creative problem-solving.

Parents and others have been fairly powerless to stand up to the nonsense of the testing interests. Until now. Fighting fire with fire, testing opponents have formed their own interest groups, including the National Center for Fair and Open Testing, based in Cambridge, Mass. After three years of fighting, they can record some progress, at least among young children: Mississippi ends state-mandated kindergarten testing. North Carolina bars testing below third grade. Arizona halts first-grade testing. The National Association for the Education of Young Children, the largest group of early childhood educators, now opposes large-scale pencil and paper standardized testing of children below third grade. So does the National Association of Early Childhood Specialists.

Regular readers know I have written about this subject before and perhaps suspect I am a nut on it. I am that, because tormenting children, physically, emotionally or mentally, is wrong. And when the schools do it, it's the most vile kind of crime.

Dennis Byrne, whose column appears Wednesdays and Fridays, is a member of the Chicago Sun-Times Editorial Board.

BEST COPY AVAILABLE

Subscribe to FairTest Publications

FairTest publications will keep you up-to-date on the growing movement for fair & open testing. You will be able to do more than just read about FairTest campaigns—you can help make them successful!

_____ *FairTest Examiner*, published quarterly...Individuals $15/year; Institutional checks $25

_____ *Fallout from the Testing Explosion: How 100 Million Standardized Exams Undermine Equity and Excellence in America's Public Schools*, by Noe Medina and Monty Neill with FairTest staff (45 pages)...$8.95

_____ *Standardized Testing Reform Sourcebook, 1988* , (200 pages)...$25.

_____ *Sex Bias in College Admissions Tests: Why Women Lose Out*, by Phyllis Rosser and FairTest staff, 2nd edition (40 pages)...$7.95

_____ *The Reign of ETS*, by Allan Nairn and Ralph Nader (550 pages)...$30

_____ *Beyond Standardized Tests: Admissions Alternatives that Work*, by Amy Allina with FairTest staff (15 pages)...$5.50

_____ *None of the Above: Behind the Myth of Scholastic Aptitude*, by David Owen (327 pp. lists at $7.95)...$5.95.

Updates, issued bi-annually (spring & fall). $7/year/topic:

- _____ *Testing Reform Initiatives*
- _____ *K-12th Grade Testing*
- _____ *University Admissions Tests*
- _____ *Teacher Testing*
- _____ *Employment Testing*
- _____ *Sex Bias in Testing*
- _____ *Test Coaching*

Order total _____

Tax-deductible contribution _____

TOTAL ENCLOSED _____
(enclose check to FairTest)

_____ Back issues of *FairTest Examiner* $4 each: Vol___, No.___ (all 6 for $20.00)

_____**Moving.** I have moved/will be moving. My new address is (Please enclose old label or write old address.):

Name ______________________ Title ______________________

Organization ______________________ Phone No. ______________________

Address ______________________ Zip ______________________

Make checks payable to FairTest and mail to: FairTest, P.O. Box 1272, Harvard Square Station, Cambridge, MA 02238

ETS to Drop NTE in Face of Rising Criticism

With a flurry of publicity, the Educational Testing Service has announced plans to replace its widely-criticized National Teacher's Exam (NTE) with a new test by 1992. While ETS's public relations staff tried to paint the decision as a natural evolution to "a new kind of assessment," the withdrawal of the NTE is really a major victory for test critics.

Across the country, a growing number of civil rights, student, and education groups have focused on the NTE as a profoundly biased and inaccurate test. In Rhode Island, an official state commission has begun investigating the test. New York legislators were scheduled to consider a proposal to apply Truth-in-Testing and item analysis provisions to the NTE later this year. Legal action on the NTE and derivative products like the Pre-Professional Skills Test (PPST) is also under consideration in Texas, Maryland and other jurisdictions. Even ETS' own 1987 Visiting Committee has criticized the test (see story, page 13).

Sample Questions

Given this mounting pressure, it's not hard to understand ETS's decision. A quick look at the only NTE made available for public inspection confirms the critics' arguments. The following items are from the November, 1982 "General Knowledge" test:

Social Studies

15. The growth of which of the following was LEAST stimulated by the growth of United States railroad lines during the nineteenth century

a) Wheat production
b) Insurance companies
c) Steel production
d) Iron mining
e) Tobacco production

Comments: Correct answer (e) was selected by 15% of the examinees. The answer requires either fantastically detailed knowledge of nineteenth century U.S. economics, or perhaps a lucky guess based on knowledge of main railroad locations. In any event, is this information even valuable for a U.S. history teacher to have in his or her head?

Mathematics

7. Given:

(1) All education majors teach.
(2) Some education majors have double majors.
(3) Some mathematics students are education majors.

Which of the following conclusions necessarily follows from (1), (2) and (3) above?

a) Some mathematics students have double majors.
b) Some of those with double majors student teach.
c) All student teachers are education majors.
d) All of those with double majors student teach.
e) Not all mathematics students are education majors.

Comment: Correct answer (b) was selected by only 11% of the examinees, perhaps because the answer is in fact incorrect according to the rules of logic. Answer (b) requires only (1) and (2), not (3). Answer (e) follows from (3) alone. No "correct" answer requires (1), (2) *and* (3), but both (b) and (e) are correct in that there is sufficient information to choose either one, and neither is contradicted by (1), (2) or (3).

(continued on page 7)

FairTest
National Center for Fair & Open Testing
P.O. Box 1272
Harvard Square Station
Cambridge, MA 02238
(617) 864-4810

Address Correction Requested

Bulk Rate
U.S. Postage
PAID
Cambridge, MA
Permit No. 53278

Complimentary Copy.
Please subscribe today.
See page 15.

FairTest EXAMINER

The National Center for Fair & Open Testing Volume 3, Number 1 Winter 1989

NCAA Told to Justify or Drop Test Score Use in Proposition 42

FairTest has called on the National Collegiate Athletic Association (NCAA) either to make public validity studies justifying its use of SAT and ACT scores in its controversial new Proposition 42 (a more restrictive version of Proposition 48) or to drop the requirement.

"Prop. 42" is a regulation which bars universities from granting athletic scholarships to students who do not obtain both a 2.0 grade point average in a core curriculum and either a 700 on the SAT or a 15 on the ACT. Cultural bias in the tests ensures that most of the estimated 600 students per year who will be denied scholarships (and therefore, in many cases, a college education) will be Black and lower-income students who would otherwise likely perform well in college.

Photo by Tom Tyburski

Participants in news conference announcing ACLU suit against New York State. Pictured above: Sarah Stockwell of FairTest, plaintiff Khadijah Sharif, Isabel Katz Pinzler of ACLU, Rob Levy of the NY Civil Liberties Union and Sheree Lewis, another plaintiff in the suit. See story, page 3.

Under Prop. 48, a student meeting either the test score or grade point average requirement (a "partial qualifier") could be granted a scholarship but could not play freshman year. The student could play as a sophomore if his or her first year grades were acceptable. NCAA statistics show that nearly 90% of all "partial qualifiers" —students who would have been excluded had the more rigid Prop 42 been in effect

(continued on page 5)

Cinthia Schuman is New FairTest Executive Director

FairTest's year-long search for a new Executive Director has ended with great success. Cinthia Schuman, currently Program Associate at the Rockefeller Family Fund in New York, will move to Cambridge to take over the position around March 1st.

Ms. Schuman, a lawyer with experience in both civil liberties and women's issues, has long worked against bias in testing. At the Rockefeller Family Fund, she helped FairTest win its first foundation grants and has been a valued advisor over the past several years.

"It's been a natural progression," Ms. Schuman explained, "I first campaigned against testing overuse and tracking as a high school student. In college, my anger about race and gender bias grew even stronger. Then as FairTest's program officer, I became an inside advocate for its cause. Now as FairTest's Executive Director I can help lead the only national group doing testing reform work full-time."

"FairTest is at a pivotal point in its development," said Ms. Schuman. "The organiza-

(continued on page 2)

CONTENTS

FairTest Moves	2
Lawsuit Filed Against NY	3
LSAT Rescored	5
New ACT and SAT	5
Report Blasts Tests in Ark.	6
Missouri Alternatives	7
Teacher Tests Hit for Bias	8
Teacher Programs Closing?	9
Stop Testing Abuse: CEEB	11
Madaus to Study Industry	11
Alice in Wonderland?	13

TM017189

New FairTest Executive Director *(from page 1)*

tion is on a roll. Its reputation, board and direction are solid. But FairTest now must broaden its support base to obtain the resources necessary to expand and strengthen its programs. I look forward to the challenge."

Before joining the Rockefeller Family Fund staff in 1985, Ms. Schuman was a Staff Associate with the American Civil Liberties Union and Executive Director of Divorce Equity, Inc. in Cleveland, Ohio. She also served internships with the Lawyers Committee for Human Rights and the New Jersey Public Advocate. Ms. Schuman is a graduate of Oberlin College and New York University School of Law, where she won a Root-Tilden scholarship for exceptional academic ability and commitment to public service. "Cinthia Schuman was clearly the best candidate we saw from over 150 applicants for the position," explained FairTest Treasurer Bob Schaeffer, who coordinated the Executive Director search. "The Board of Directors looks forward to working closely with her as together we build up FairTest's program, staff and finances."

You can greet Cinthia Schuman at FairTest's Open House in our new Cambridge offices on March 10 (see below), at our Southern Conference in Atlanta on March 17-19 (see p. 16), or at future speaking engagements across the country.

Cinthia Schuman

FairTest Moves to New Office: 342 Broadway, Cambridge, MA 02139

On December 3, 1988, FairTest moved its office from Harvard Square to a spacious new location at 342 Broadway, Cambridge, Massachusetts 02139. The telephone number remains the same: (617) 864-4810. Please be sure to address all future correspondence, mailings, etc. to this new address.

Our old office, while in a prime location, was cramped and nearly bursting at the seams. The new space, an 1840's house renovated by the live-in owners, is large enough for our current full-time staff of five, and will give us room to grow as our staff and programs expand.

Open House

On March 10, 1989 from 5-7PM FairTest will hold an Open House to enable our friends and supporters to meet the staff and new Executive Director. All are welcome.

The National Center for Fair & Open Testing (FairTest)
342 Broadway, Cambridge, MA 02139; telephone (617) 864-4810.

Staff: Mike Friedman, Publications Manager; Mary Gilleece, Office Manager; Noe Medina, K12 Consultant; Monty Neill, Managing Director; Sarah Stockwell, Admissions Test Coordinator; Bob Schaeffer, Communications.

Board of Directors: Prof. Denise Carty-Bennia, Northeastern University Law School, Chair; Dr. Pamela George, NC Central University; Paul Pottinger, New Manhattan Corp.; William Robinson, Dean, D.C. School of Law; Bob Schaeffer, Treasurer; Steve Solomon, Esq.; Chuck Stone, *Philadelphia Daily News*; Ron Vera, Esq., Barbosa & Vera; John Weiss, President.

Partial funding for FairTest's work comes from the Ford Foundation and the Rockefeller Family Fund.

Correction

In the Fall 1988 *FairTest Examiner* article "ETS to Drop NTE," Dr. Lorrie Shepard was erroneously cited as the author of the study finding that 40% of the questions on one form of the NTE Test of Professional Knowledge had either no correct answer or more than one correct answer. In fact, Dr. Linda Darling-Hammond of the Rand Corporation reached this conclusion in "Teacher Knowledge: How Do We Test It?" (*American Educator*, Fall 1986). Unfortunately, this error was caught after the issue went to press. FairTest apologizes to both fine scholars.

University Admissions Testing

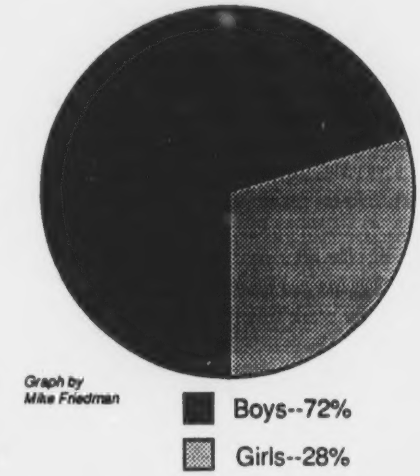
Lawsuit Claims Sex Bias in $40 Million New York Scholarships

A coalition of civil rights and women's groups has sued New York State claiming sex discrimination in the annual award of up to $40 million in Regents and Empire Excellence Scholarships solely on the basis of SAT or ACT scores.

FairTest worked closely with the Women's Rights Project of the American Civil Liberties Union (ACLU) and the New York Civil Liberties Union (NYCLU), to prepare the case. The suit was brought in federal court on behalf of ten female high school seniors, the National Organization for Women, New York State (NOW-NYS), and the Girl's Clubs of America, Inc. At a New York City news conference announcing the suit, Isabelle Katz Pinzler, Director of the ACLU Women's Rights Project and lead counsel in the case, said, "The SAT discriminates against girls and underpredicts their college performance. The State knows this. Yet they continue to use it as the only criterion for awarding state scholarships."

Research conducted by the New York Public Interest Research Group and FairTest and confirmed by the New York State Department of Education revealed that through 1987, boys received nearly three-quarters of top New York State scholarships awarded solely on the basis of test scores. For 1988, the New York legislature changed the award formula to include high school grades. As a result, girls won a fairer share of awards. Because the reform was not renewed, scholarships are again slated to be awarded solely on the basis of test scores.

Projected Empire Awards of Excellence Using "Test Scores Only" Formula

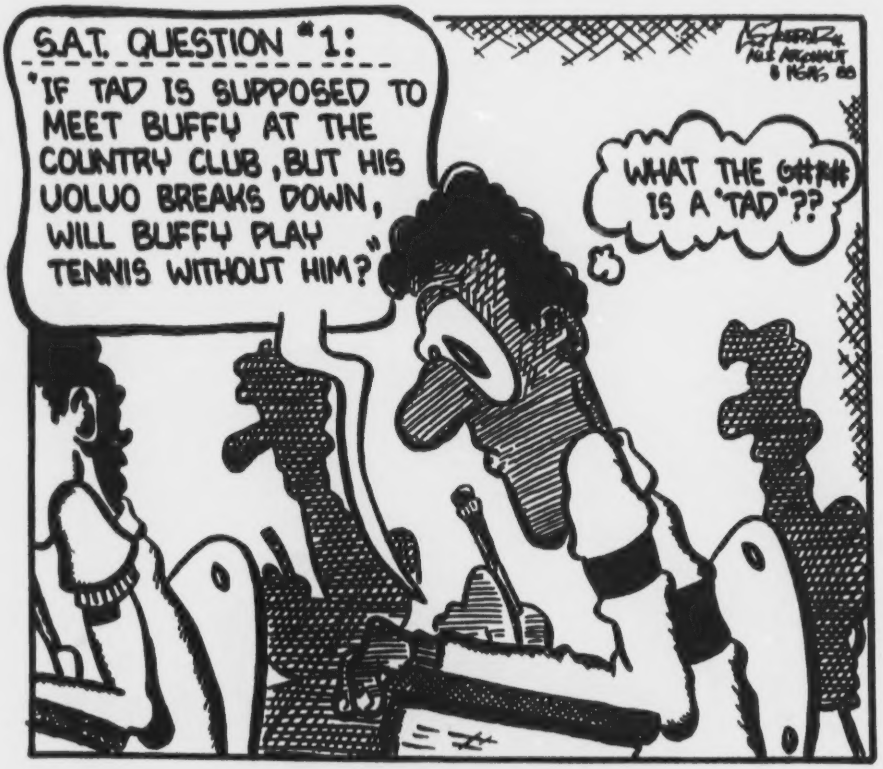


Graph by Mike Friedman

Robert Levy, NYCLU staff counsel explained, "The issue here is simple. By returning to a formula which discriminates against women, New York is consciously denying young women a fair opportunity to compete for scholarships. That violates equal protection of the law."

The plaintiffs have asked a Federal District Court to enjoin New York state from using the SAT-only formula for this year, citing irreparable harm to females who are eligible for the scholarships. The complaint suggests that, instead, grade point average should be part of the formula, since last year's formula of 50% GPA and 50% test scores produced much fairer results. Causes of action are based on Title IX regulations and the due process and equal protection clauses of the Fourteenth Amendment to the U.S. Constitution.

The issues in this case are similar to those raised by FairTest and a coalition of 50 civil rights, education reform, feminist, and consumer groups who have called on the National Merit Scholarship Corporation to stop using PSAT/NMSQT test scores as the sole qualifying criterion for its awards.


S.A.T. QUESTION #1:
"IF TAD IS SUPPOSED TO MEET BUFFY AT THE COUNTRY CLUB, BUT HIS VOLVO BREAKS DOWN, WILL BUFFY PLAY TENNIS WITHOUT HIM?"
WHAT THE G#%# IS A 'TAD'??

BEST COPY AVAILABLE

University Admissions Testing

New ACT and SAT: Cosmetic Changes Mask Fundamental Flaws

Both the American College Testing Program (ACT) and the Educational Testing Service (ETS) have recently announced changes in their university admissions tests. Next fall ACT will begin offering a significantly revised ACT Assessment. Apparently caught off-guard by ACT's initiative, ETS announced vague plans to update the SAT as well.

While the changes, at least in the case of the ACT, may appear substantial, in both cases they are likely to be little more than cosmetic. Neither "new" version will address the exams' fundamental flaws: unfair biases against women and minorities and very limited predictive validity. Almost certainly the tests will continue to measure little more than the ability to take standardized tests.

ACT claims that the content changes will shift its emphasis away from simple knowledge of facts to higher-order reasoning skills. Other alterations include an increase in subscores from four to twelve and the elimination of social studies and natural science sections. The changes constitute the first major overhaul of the ACT in its 30-year history.

The only set of questions from the new ACT released thus far reveal the on-going problems with the test. According to ACT, one improvement is that all the information the test-taker needs to answer the question will be provided. However, in the examples released by ACT, *too much* information was provided: the entire introductory paragraph of the science-oriented set of questions was unneccessary for solving the problems, which required interpreting a table of data. This is an example of a trick question which measures test-wiseness as much as the ability to deal with a scientific table.

ACT AVERAGE COMPOSITE SCORES
BY GENDER AND RACE (1987)
(1 to 36 point scale)

White M = 20.4
Asian M = 20.3
Asian F = 19.3
White F = 18.9
Puerto M = 17.8
Mexican M = 16.2
Puerto F = 16.1
Native M = 15.3
Mexican F = 14.8
Native F = 14.1
Black M = 13.8
Black F = 13.2

Asian = Asian-Americans
Puerto = Puerto Ricans/Cubans and other Hispanics
Native = Native Americans/Alaskan Natives
Mexican = Mexican Americans/Chicanos

(Source: *Special ACT Assessment Reference Norms*, National 1986-87.)

ACT has also announced plans to market ACT-like tests for 8th graders, 10th graders and college sophomores.

After the SAT's rival unveiled its new test, ETS and the College Board announced that they had begun a three-year review of their own university admissions test. Though no details were provided, the groups said the process could lead to the inclusion of non-multiple-choice questions, writing samples and open-ended questions to help make the SAT less coachable. A score in writing and an additional math score might also be added to the current Verbal and Math scores. However, the project manager for the new test admitted that, in the end, the SAT may not be changed at all.

The move toward a revised SAT appears to be motivated as much by marketing considerations and an attempt to capitalize on the growing misuse of test scores as by a genuine desire to create a better exam. College Board President Donald Stewart, in explaining reasons for the changes in the test, told the *New York Times*, "People are increasingly using the SAT to place students in remedial or advanced courses, award scholarships, evaluate high school curriculums and compare schools at the local, state and national levels." But all of these uses are now discouraged in the College Board's *Test Use Guidelines*, and none can be resolved by simply revising the test.

High sounding statements and cosmetic changes are insufficient: the College Board and ETS have an obligation to ensure their products are properly used.

University Admissions Testing

NCAA Proposition 42 *(from page 1)*

—have been Black (see Fall, 1988 *Examiner*).

In a letter to NCAA Executive Director Dick Schultz, FairTest said "By adopting Prop. 42 the NCAA is asserting that student athletes who score less than 700 on the SAT or 15 on the ACT cannot do college level work. Guidelines for proper test use mandate an up-to-date validity study proving that such a conclusion is accurate before any test score requirement can be imposed. But, evidence already exists that the NCAA is wrong. Recent reports demonstrate that minority athletes with low scores can do satisfactory work in college and complete academic degree requirements."

"Unless the NCAA can produce validity studies justifying its position," FairTest concluded, "Proposition 42 is a blatantly improper use of test scores which must immediately be abandoned."

FairTest cited current College Board *Guidelines on the Uses of Test Scores and Related Data* discouraging the use of cut off scores "unless properly validated" and similar statements from the American College Testing Program. College Board guidelines also urge users of SAT scores to "validate data used in the selection process regularly (generally at least once every three years) to ensure their continuing relevance . . ." and to "take into appropriate consideration predictions of performance for applicant subgroups—men and women, ethnic groups, international students, adults and those interested in different academic programs— in developing equitable admissions policies and practices."

FairTest has also written to the College Board and the American College Testing Program calling on them to enforce their guidelines by withholding test scores from schools which use them in connection with Proposition 42. "Admissions test manufacturers have an obligation to ensure their products are properly used," FairTest's letter said. The proof of their sincerity comes in enforcement."

The lack of demonstrated validity, the prevalence of bias and the misuse of test scores in a manner which excludes low-scoring athletes has aroused substantial controversy. John Thompson, Georgetown University and 1988 U.S. Olympic basketball coach, attracted attention to the issue on January 14 by walking off the court before a game with Boston College. Thompson denounced the denial of scholarships, "[Athletes] will no longer have the opportunity to show that a poor test score... is not a result of the lack of native intelligence, but is a result of the cultural bias of the test and the deprivation that has existed in their lives because of socieconomic and racial issues that are, unfortunately, inherent in our society."

> **"A poor test score...is not a result of a lack of native intelligence, but a result of the cultural bias of the test."**
> —Georgetown Basketball Coach John Thompson

Boston University coach Mike Jarvis agreed, pointing out that "So much of the SATs is based on what you've been exposed to. But exposure is not the same as intelligence."

Testing reform advocates hope that the controversy around Prop. 42 will expand to include the role of test scores in blocking equal educational opportunity for minorities, women, and low-income students in general, not just athletes.

LSAT Rescored

After a test-taker repeatedly challenged one item on recent Law School Admissions Tests (LSAT), Law School Admission Services (LSAS), which administers the exam, withdrew the question, rescored some 23,000 tests and raised nearly 750 scores.

The question was #13 on the June and October 1988 LSAT's:

> "The average electric utility plant can generate one kilowatt of electricity at a cost of six cents. However, whenever a company adds a nuclear power plant, the average cost per kilowatt generated rises dramatically. This proves that nuclear power is not cost efficient.
>
> 13. Which of the following beliefs is necessary for the author's conclusion to be reasonable?
>
> (A) Nuclear power is dangerous.
> (B) Coal is dirty.
> (C) Oil is in short supply.
> (D) Utilities are poorly managed.
> (E) Cost is an important determinant in choosing a power generation source.
>
> Desired response: (E).
>
> None of the beliefs are necessary to answer the question!

SAT Bibliography

FairTest has produced a new bibliography listing selected items on gender and race discrimination, coaching and misuse of the Scholastic Aptitude Test. For a copy of the *Selected Bibliography on the SAT: Bias and Misuse*, please send your request with a stamped self-addressed envelope to FairTest, 342 Broadway, Cambridge, MA 02139.

Grade School Testing

Arkansas Report Criticizes Testing

"Teaching to the MPT (Minimum Performance Test) is a large share of the curriculum in many settings....and tend(s) to distort the educational process itself," concludes *Fueling the Promises of School Reform*, a 1988 survey on school reform in Arkansas issued by the Winthrop Rockefeller Foundation.

The report interviewed faculty in 15 schools to determine the impact of new state requirements put into effect as part of a school reform initiative. Explained one typical teacher, "The pressure of test scores should be reduced. Test scores are a method of measurement, but their importance has been blown all out of proportion."

On other testing issues, the report said that teachers believed they can identify students in need of remediation without the test; tests ought not to be the primary vehicle for accountability; and Blacks on average score less well than Whites on the MPT.

Also, teachers regarded teacher testing as the least effective reform instituted in Arkansas and blamed it for a sharp decline in teacher morale. Teachers also criticized increased centralization and paper work, some of which was associated with testing.

To date, use of the MPT apparently has not caused an increase in the dropout rate or in special education placements (which, however have not decreased). But only in 1988 did the MPT become "a definitive high stakes test," so both rates could increase as tests increasingly dominate decision-making.

> **"Teaching to the MPT is a large share of the curriculum in many settings....and tend(s) to distort the educational process itself"**
>
> —Winthrop Rockefeller Foundation Report on Arkansas Schools

(*Fueling the Promise of School Reform* is available from the Winthrop Rock Foundation, 308 East Eighth St., Little Rock, AR 72202.)

For more information on K-12 test misuse, order FairTest's report, *Fallout from the Testing Explosion*. See page 15 for ordering information.

Rethinking Schools

Rethinking Schools, a Milwaukee-based progressive school reform magazine has just released its January/February 1989 issue, focusing solely on standardized testing.

Topics covered include:

—The case against standardized achievement tests.

—NAEYC's position paper on testing.

—Linda Darling Hammond on "Alice in Testing Land"

—Alternative assessment in high schools.

—Bibliography on rethinking assessment.

—Testing in Milwaukee, and much more.

To order a copy of *Rethinking Schools*, send $2.00 ($10 for one year) to P.O. Box 93371, Milwaukee, Wis. 53202.

K12 In Brief

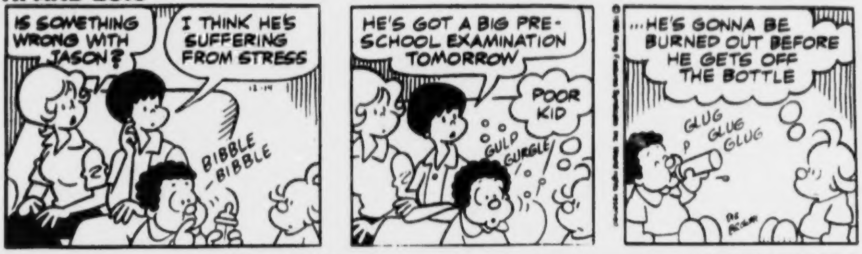
All Still Above Average

Further proof that standardized achievement test scores are inflated was provided when the Southern Regional Education Board (SREB) released its finding that 90% of southern students' scores on these exams were at or above national averages. Mark Musick of the SREB told *Education Week* (10/26/88), "We do believe this raises legitimate questions in the public's mind about what the scores mean." The SREB study found both an increase in the use of standardized testing in the south and an increasing effort to reduce test use. The report, "Measuring Student Learning," is available from the SREB for $6 (592 10th St. NW, Atlanta, GA 30318).

Coaching Works for Youngsters Too

According to a study reported by Barbara Keough of UCLA, teaching young children testing skills totally devoid of content can lead to substantial score increases. Working with a group of Hispanic children, she reported, "We trained one group of youngsters in testing experiences. The training was totally non-content. We got a whopping main effect. When the children who had been trained to take tests were tested they did extraordinarily well...Such findings raise issues regarding what's being assessed, and how much of early screening is really testing a child's aptitude and how much is an indication of school experiences related to taking tests" (*Communique*, December 1988).

HI AND LOIS


IS SOMETHING WRONG WITH JASON?
I THINK HE'S SUFFERING FROM STRESS
BIBBLE BIBBLE
HE'S GOT A BIG PRE-SCHOOL EXAMINATION TOMORROW
POOR KID
GULP
GURGLE
...HE'S GONNA BE BURNED OUT BEFORE HE GETS OFF THE BOTTLE
GLUG GLUG GLUG

Grade School Testing

Missouri Developing Alternatives to Standardized Testing

The Missouri Department of Elementary and Secondary Education is developing alternatives to standardized testing for use in pre-school, kindergarten and grade 1 programs. The alternative assessments are being tested in 10 pilot programs and will be made available statewide in 1990. Missouri does not mandate standardized testing prior to grade two.

The alternative assessments will flow from newly-established goals and reinforce them. According to Debbie Murphy, Director of Early Childhood Education, this should make the new evaluations more useful for teachers than standardized tests, which have become separated from classroom experience.

The Department is also attempting to identify developmental patterns across cultures. For example, how do children from a variety of backgrounds approach spelling at different ages? To avoid emphasizing competencies that are culturally specific, the samples include a mix of children. The evaluations will be designed to encourage diversity in learning styles and skills, not just those strongly associated with the white middle class.

Missouri is using teachers as the experts to create the new assessments. "It's time to stop blaming teachers," explained Murphy. "We have to help them become autonomous professionals and give them power." Thus, teachers are involved in creating and critiquing the alternatives that are developed.

In the Missouri model, the first use of the new assessment will be to help the child by, for example, diagnosing a problem in emerging literacy. Guides for teacher observation of student processes and products in which the teacher is in control of the assessment will be provided. Rather than one 'correct' response from either student or teacher, the model will provide a list of strategies and approaches to diagnosis and, thus, to assistance.

The assessments will be intertwined with real classroom activities. For example, teachers will be given guides for observing how children act in a card game: do they hide their cards, take turns, show numerical ability? For evaluating a whole class, the assessments will necessarily be less detailed. Instead, they will provide a "mental scaffolding" for understanding the learning that has occurred. Teachers will accumulate substantial raw data which they can use in the class. This data can be summarized descriptively for school or system use and will also provide quantifiable information about class and grade achievement.

"Tests," Murphy remarked, "mostly sit in a file drawer. That's the best thing that can happen to them." But, she cautioned, developing alternatives is not easy. "We don't have all the answers. It's scary trying to help teachers while making sure the new assessments don't harm children or teachers." For example, the Department has not yet considered how to ensure that individual teachers do not subvert the new procedures with race or class biases.

By 1990, the Department plans to make the alternatives available to teachers who want to use them, but will not mandate that teachers must use them. The Department will offer extensive in-service training to aid teachers in the use of the new assessments. The Department also plans to work with administrators, and with parents so they will understand and support the new methods.

"Wake up! Time to start thinking about the SAT."

Teacher Testing

Tests Keep Thousands

Nearly 38,000 Black, Latino, American Indian and other minority teacher candidates have been barred from elementary and secondary classrooms by teacher competency tests in the past few years. This is the startling conclusion of *The Effects of Competency Testing on the Supply of Minority Teachers*, [illegible] Dr. G. Pritchy Smith, Professor at the University of North Florida.

According to Smith, the actual number of minorities eliminated by tests like the National Teachers Exam is probably even higher because many states do not report data by race/ethnicity or data for all test administrations. His analysis of 19 of the states now requiring written exams found that minority pass rates were consistently lower than White pass rates. In 16 of the states, Black pass rates were below 50%.

Smith's report proves that teacher testing is a major factor in the likely reduction of the minority portion of the nation's teaching force from 12% to 5% by the year 2000, when over 1/3 of the nation's students will be from minority groups.

Florida Case Study

In a detailed study of Florida, Smith found that testing excluded Blacks from upward mobility: From grade school to high school graduation and entry into public universities and teacher education programs, a series of standardized, multiple-choice exams excludes significantly more Blacks and Latinos than Whites.

For example, in Florida, SAT or ACT test scores above a specified level are required for admission into teacher training programs. This requirement eliminated 80% of the Black and 61% of the Hispanic candidates, but only 37% of the Whites. In 1985-86, only 98 Black students were admitted into the state university system teacher education programs. But performing satisfactorily in a program is not enough to become a classroom teacher. Another test must be passed. On the Florida Teacher Certification Examination, Black pass rates are 37%, Hispanic 53% and White 89%.

As a result, only about 34 Blacks, out of an age cohort of 25-30,000 Blacks, will graduate from a Florida public university with a teaching certificate. Just 170 more will graduate from private universities and colleges. Thus, testing effectively erases minorities from the teaching profession.

Closing College Programs

Nationally, not only are tens of thousands of minority applicants kept from teaching, but many predominantly Black or Hispanic teacher education programs face the threat of being closed down due to low pass rates on the teacher tests. That would further reduce the number of minority teachers. Yet, Smith noted, such sanctions already have been imposed on historically Black colleges (HBCs) in Florida. Similar proposals to close programs are under consideration in Alabama and Arkansas.

Setting Cut Scores

Smith observes that each state sets the cut scores to be used on a test. For example, the passing score on the NTE, the mostly widely used test, varies from state to state. Typically, a cut-score is determined by a committee which assesses the percentage of examinees they think will answer each question correctly and then derives a passing score. But, as Terry Haertel of Stanford has explained, it is not possible to set a sole cut-score "capable of predicting adequate versus inadequate performance in diverse teaching settings" (*Chronicle of Higher Education*, Nov. 30, 1988). Yet in a study of NTE use in North Carolina, Jim Gray of the Lawyer's Committee for Civil Rights Under Law observed that 46% of all test takers had scores within 10 points in either direction of the cut score. Often, Black scores cluster just below the cut-off level. Thus, a question or two on the test can make a substantial difference in the number of minorities certified to become teachers.

Prof. G. Pritchy Smith of the University of North Florida

Valid Instruments

Smith's report demonstrates in detail the devastating impact tests are having on minority teachers. If the exams also lack validity, then there can be no excuse for their continued use.

Haertel also argued that the tests lack validity since the correlation between test score and teacher performance is about zero. Even though Blacks who do pass have lower average test scores than Whites who pass, "the only study of beginning teacher programs...that reports data by race, black teachers out-performed white teachers on beginning teacher classroom performance evaluations."

Not only do the tests not predict whether those who pass will be good teachers, but no evidence exists to show that those who fail the test would be inadequate teachers. Explained Haertel, "(I)f test scores cannot be shown to predict classroom performance, then rational standard-setting is impossible."

Teacher Testing

of Minorities Out of Teaching

Another crucial point not examined in Smith's article is that of racial bias in the construction of the exam. A study by ETS's Medley and Quirk published in 1974 indicated that an equally difficult NTE General Knowledge exam with different content could have equal pass rates for Black and White applicants.

Thus, these tests are being used despite a lack of demonstrated predictive validity, clear racial bias, and cut scores that are set under a process that amounts to guessing. As a result, tens of thousands of minorities are being excluded from teaching. There appears to be no justification for any state requiring their use as a condition for obtaining teaching licensure at least until the new generation of tests is available. And any new instruments must meet adequate standards of validity.

Pritchy Smith's report was prepared for the National Education Association and the Council of Chief State School Officers and was part of a series on teacher testing. Because his report still had not been made public one year after he completed it, Smith released it himself. Among the other unreleased reports is one addressing the question of teacher test validity.

NTE Question Further Flawed

Examiner reader Adam Stein of Chapel Hill, NC wrote us recently, pointing out that one of the flawed NTE questions we published in the Fall, 1988 issue (p.16) was impossible to answer, using the items given in the question.

The logic question, from the November, 1982 NTE General Knowledge test, is reprinted at the right, along with FairTest's comments.

Stein pointed out that since none of the given statements include any evidence about *student* teaching, "to conclude that (b) is correct requires an inference that the education majors are student teaching, as contrasted with some other kind of teaching. That inference is certainly reasonable. However that inference does not 'necessarily follow' from the given."

Indeed, Adam Stein is perfectly correct. This simply reinforces FairTest's contention that the NTE is a poorly-worded, flawed exam that fails to measure anything remotely resembling classroom performance of prospective teachers. Any professional logician who wishes to comment further on this question should write to Mike Friedman at FairTest.

7. Given:

(1) All education majors teach.

(2) Some education majors have double majors.

(3) Some mathematics students are education majors.

Which of the following conclusions necessarily follows from (1), (2) and (3) above?

(a) Some mathematics students have double majors.

(b) Some of those with double majors student teach.

(c) All student teachers are education majors.

(d) All of those with double majors student teach.

(e) Not all mathematics students are education majors.

Comment: Correct answer (b) was selected by only 11% of the examinees, perhaps because the answer is in fact incorrect according to the rules of logic. Answer (b) requires only (1) and (2), not (3). Answer (A) requires (2) and (3), but not (1). If (B) is "better," it would require an interpretation of the phrase "some of those" to somehow exclude "those" covered under given (2).

Quote of the Month

"If you want to change what goes on in schools, you must change the way things are assessed. As long as we use standardized, multiple-choice, machine-scored tests, that's going to dominate what goes on."

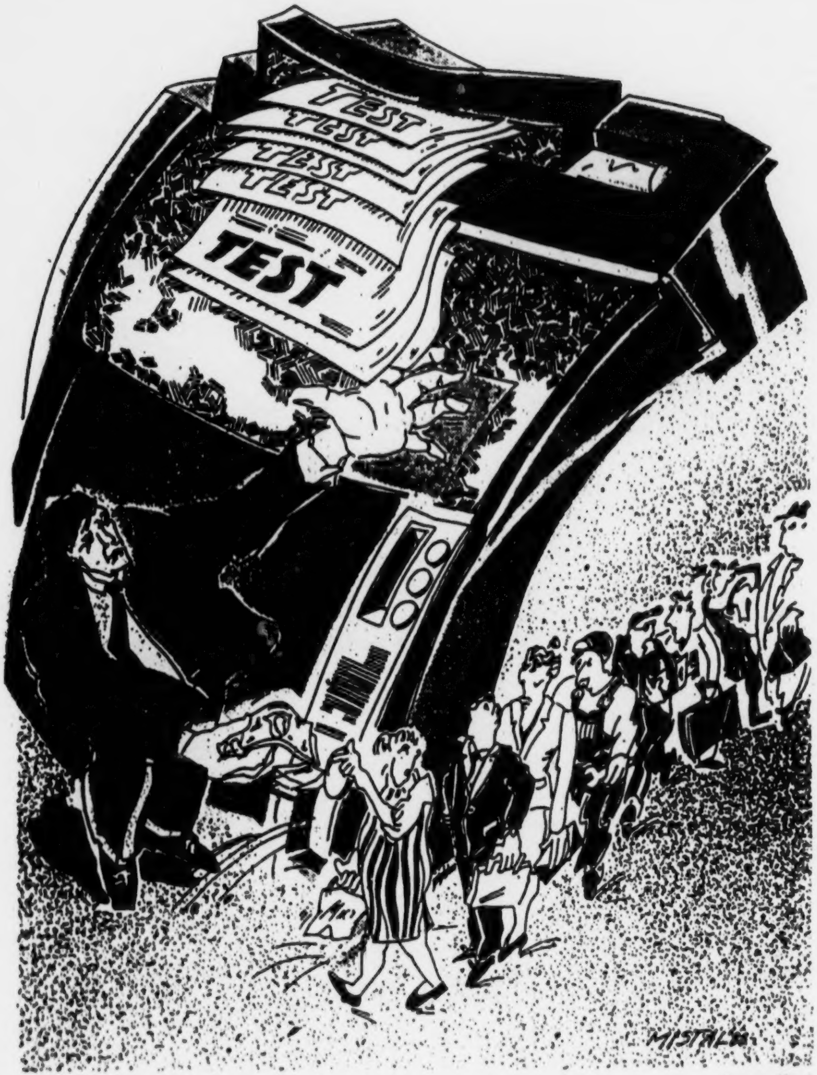
—Howard Gardner,
Harvard School of Education

Black Teacher Programs to Close?

Unless their students score higher on the NTE, teacher education programs at 10 of North Carolina's 43 colleges and universities that train teachers will be closed as early as 1990.

All five predominantly Black schools of education in the UNC system will be closed unless 70% of the school's graduates pass the NTE. The other five schools at risk of losing state accreditation are private, predominantly Black schools. In the past two years, fewer than 60% of Black test takers passed the test.

Other states have also threatened to close teacher programs whose graduates do not attain sufficiently high pass rates on teacher tests. However, no study has been done to validate the use of the NTE or similar instrument as a device to measure school quality.



Catherine Mistal

Warfare Breaks Out Between Prep School Testing Companies

After three decades of enjoying a monopoly position in the independent (private and parochial) school admissions testing process, the Secondary School Admissions Test (SSAT) suddenly has a competitor. In September 1989, the Education Records Bureau (ERB), which previously handled administration of the SSAT on a subcontract basis, will begin offering its own Independent School Entrance Examination (ISEE).

The result is a marketing war between the two companies with each promising "new, improved" features, but neither dealing with the old flaws of standardized admissions exams.

Rumors abound about ERB's reason for going up against its historic ally. Finances, egos, and even a desire to make a better test have all been mentioned. But whatever the motivation, ERB has already launched a massive public relations campaign to get schools to use its new product. Among ISEE's marketing claims are more in-depth information about students and delivery of score reports to schools anywhere in the world within 72 hours included in the $40.00 fee.

The Secondary School Admissions Test Board is not sitting still. It has countered with what it calls "The New SSAT" which, not surprisingly, also promises more detailed student profiles. In addition the SSAT will now attempt to predict SAT score ranges.

For students and their parents, all this means that there will be two tests to choose between. But there's little evidence that either the "new" SSAT or the "new" ISEE meets minimum standards for accuracy, fairness or openness. In place of marketing gimmicks, the public deserves honest and independent data evaluating both exams. Extending Truth-in-Testing, Item Analysis and Bias Reduction requirements to all secondary school admissions tests, no matter how new they claim to be, would be the first step.

BEST COPY AVAILABLE

Madaus Receives Grant to Study Testing Industry

Dr. George Madaus, director of the Center for the Study of Testing, Evaluation and Education Policy (CSTEEP) at Boston College has recently received a grant of $197,000 from the Carnegie Corporation to investigate the feasability of organizing a new independent group to monitor and regulate so-called "high stakes" testing.

Madaus and BC colleague Dr. Walter Haney will use the grant to study the various ways in which organizations and industries are regulated or regulate themselves, and evaluate the effectiveness of the testing industry's self-regulation in reducing or eliminating bias and test misuse. In addition, they will review ETS's *Standards of Quality and Fairness*, the *Standards for Educational and Psychological Testing* developed by the American Psychological Association, and the recently released *Code of Fair Testing Practices in Education*.

Haney and Madaus then hope to assemble a group of prominent representatives of academic, political, educational and testing industry circles to form the nucleus of a Board of Directors for an independent agency to monitor the testing industry. CSTEEP has applied for a follow-up grant in order to implement the recommendations of the current study.

For more information on the study, write: Dr George Madaus, CSTEEP, Boston College, Chestnut Hill, MA 02167.

Black Police Officers File Suit Against Kansas City, KS

Nine Black Kansas City, Kansas police officers have filed a class action lawsuit alleging that the city has violated their rights under the Civil Rights Acts of 1866, 1871 and 1964 and the 13th and 14th Amendments to the Constitution, by practicing racial discrimination in promotions to sergeant, detective and lieutenant. All promotions to the upper ranks are done on the basis of scores on standardized tests.

According to the plaintiff's attorney, Blaise Plummer, the city adopted testing for police department promotions in 1977. Until that time, Plummer said, Black officers were well represented in the upper ranks of the Kansas City police department. Since the institution of standardized testing, the number of Black upper level officers has declined severely, due to retirement, attrition and non-promotion of Blacks.

In the suit, the plaintiffs allege that the city showed "the purpose and intent to discriminate" by systematically abolishing the promotion pool every time a Black officer reached the top position. This resulted in re-testing, the creation of a new pool of candidates and the relegation of Black officers to the bottom of the list. In addition, the city is accused of having only White officers in charge of training and testing, and providing more access to testing and training materials to Whites than Blacks.

The officers have requested damages "in excess of $10,000" as well as back pay and attorneys fees. Filed in December 1988, the suit, *Simmons, et al. v. City of Kansas City, Kansas*, also names Police Chief Allan Myers and City Administrator David Isabell as defendants.

Stopping Test Misuse:

Lip Service, Not Enforcement

Top officials of the College Board (CEEB) and American Council on Education (ACE) have called on school officials and the news media to stop using test scores in rating post-secondary institutions.

In a letter to universities and colleges, CEEB President Donald Stewart and ACE President Robert Atwell declared "The tests are meant to assess individual achievement and ability, not to compare the institutions students attend, or the schools, school districts and states they hail from."

But the strong statement has done nothing to change the behavior of the College Board, which can control the use of SAT scores because it owns the test. Dr. Stewart, has, for example, done nothing visible to block the use of improper comparisons based on SAT scores when they appear in newspapers and magazines.

Nor has the College Board taken measures to stop the rampant use of cutoff scores in admissions decisions, though Dr. Stewart also said "The SAT score is but one factor that can be taken into consideration in making admissions decisions."

More useful than hollow statements would be for the College Board to declare that it will withhold test scores from institutions which misuse the results, and that it will employ its considerable public relations muscle in censuring individual institutions. That would be a more effective way to permanently stop the misuses the College Board and ACE so eloquently decry.

"Stand and Deliver"

On March 15, 1989, the Public Broadcasting System's "American Playhouse"will show the hit film "Stand and Deliver."

"Stand and Deliver" is based on the story of Jaimie Escalante, played by Edward James Olmos, a math teacher at Garfield High School in Los Angeles, and his attempts to teach advanced calculus to a group of poor, disadvantaged students.

The movie chronicles their success and their struggle against the Educational Testing Service, which questioned their scores on an Advanced Placement examination.

Check your local listings to confirm the time and channel.

The New York Times

Copyright © 1988 The New York Times

NEW YORK, WEDNESDAY, DECEMBER 7, 1988

About Education

Fred M. Hechinger

Choose one: Standardized tests are a) gaining favor; b) losing favor; c) neither. (Answer: b)

Revolutions can succeed only if there is enough dissatisfaction with things as they are and if there is a banner around which to rally. A growing number of education experts say both of these prerequisites are in place for a revolution in testing in America.

Take first the dissatisfaction with the status quo. There is a clamor against standardized tests. Paradoxically, tests have proliferated, especially since 1983, when the report, "A Nation at Risk," by the National Commission on Excellence in Education, called for tougher academic standards. The money and time spent on tests, the toll in anxiety and the doubts about the tests' validity have left many parents and educators up in arms.

The test makers themselves have begun to issue warnings about the misuse of tests. Scholastic Aptitude Test scores are being used to rate the quality of colleges. In an open letter to educators last month, Robert H. Atwell, president of the American Council on Education, and Donald Stewart, president of the College Board, which administers the S.A.T.'s, warned that students' scores did not "necessarily say anything useful about the quality of what goes on in a college or university."

•

Aptitude tests are under fire, and with them all tests based largely on multiple-choice questions.

Much of the criticism is aimed at typical elementary-school reading tests. The critics say children's reading ability cannot be measured by having them fill in missing letters, words or sentences. The only way to judge reading, they say, is by hearing children read and talk about what they have read.

Worse, the critics add, if children, teachers and schools are judged on the basis of how well children fill in blanks, teachers will teach them to fill in blanks instead of teaching them to read and to enjoy what they are reading. Others question whether a high score on an S.A.T. vocabulary section tells much about a student's writing skill. In fact, there is much evidence that high school English teachers try to to prepare students for the S.A.T.'s by encouraging them to use big words. The result is often bad writing.

In fairness, standardized tests were not a devilish creation by enemies of good education. When the S.A.T.'s were introduced in 1926 as a major consideration in college admission, they were intended to make the process fairer. Discrimination was rampant against Jews, Catholics, blacks and women. Standardized tests were to make the process objective, and in large measure they did.

But the years brought dramatic social changes, and standardized tests have not kept pace. The pool of test-takers not only grew in number but also in diversity. Efforts were made to make the tests culture-free and color-blind. Flagrantly class-bound questions were eliminated.

•

This concern is not new. In 1899, when Charles W. Eliot, Harvard's president, called for standardized admissions tests, Ethelbert D. Warfield, president of Lafayette College, representing the opposition, called standardization a step toward dehumanizing uniformity.

The debate continues. The mood now suggests a coming revolution against standardized tests. And there is a banner around which the revolution can rally. It was raised in 1986, when Lee S. Shulman, a professor of education at Stanford University, began the Teacher Assessment Project. Commissioned by the Carnegie Corporation of New York, the project is devising tests to certify teachers based on actual or simulated performance and on portfolios of past accomplishments.

When asked whether such new non-standardized tests might revolutionize all testing, including the S.A.T.'s, Professor Shulman said, "It's not too much to dream about."

Since then, the dream, and the revolution, no longer seem remote. The Educational Testing Service, which produces the S.A.T.'s and a host of other standardized tests, announced last month that by 1992 it will retire the 50-year-old, much criticized multiple-choice National Teacher Examinations, used by 30 states. It will be replaced by something that sounds like the Shulman approach.

An era is ending. Dissatisfaction with the old order of testing has prepared the ground for change. The banner of the new order is in place. Attempts at standardized packaging of everything, from political candidates to teachers to college applicants to kindergarten youngsters, seem to have led to a counter-revolution against standardized testing. Introduced long ago in the interest of objectivity, standardized tests have become symbolic of a society afraid to make human judgments.

Ten Commandments for Taking a Standardized Test

by Edwin F. Taylor

I. Thou shalt not make subtle distinctions (the Universe must fit into four or five boxes.)

II. Thou shalt not think broadly. (Only one given answer can be correct.)

III. Thou shalt not have thine own ideas. (Adopt the fantasy world of the test maker or else fail.)

IV. Thou shalt not devise unusual relations among things. (A is to B as C is to what Somebody Else has in mind.)

V. Thou shalt not be different in unique ways. (You are being compared with others along a single line according to a "normal" curve.)

VI. Thou shalt be better at naming a concept than at using the concept. (Word recognition is the easiest skill to test.)

VII. Thou shalt not try too hard. (Let your middle-class background guide your choice of answer; if you are not middle class, too bad.)

VIII. Thou shalt not be confused by misleading or crowded graphic layout.

IX. Thou shalt not confuse the answer with the name of the answer, with the number or the letter of the answer, or with the box to be shaded to tell the computer the number or letter that corresponds to the answer thou hast chosen.

X. Thou shalt not be frightened.

"...don't be nervous [said the King], or I'll have you executed on the spot!")

The Original Looking Glass Achievement Test

(courtesy Lewis Carroll and Edwin F. Taylor)

"Can you do addition?" the White Queen asked. "What's one and one and one and one and one and one and one and one and one and one?"

"I don't know," said Alice. "I lost count."

"She can't do addition," the Red Queen interrupted. "Can you do subtraction? Take nine from eight."

"Nine from eight, I can't, you know." Alice replied readily, "but—"

"She can't do subtraction," said the White Queen. "Can you do division? Divide a loaf by a knife—what's the answer to that?"

"I suppose—" Alice was beginning, but the Red Queen answered for her.

"Bread and Butter, of course. Try another Subtraction sum. Take a bone from a dog, what remains?"

Alice considered. "The bone wouldn't remain, of course, if I took it—and the dog wouldn't remain; it would bite me and I'm sure *I* shouldn't remain!"

"Then you think nothing would remain?" said the Red Queen.

"I think that's the answer."

"Wrong, as usual," said the Red Queen. "The dog's temper would remain."

"But I don't see how—"

"Why, look here!" the Red Queen cried. "The dog would lose its temper, wouldn't it?"

"I suppose it would," Alice replied cautiously.

"Then if the dog went away, its temper would remain!" the Queen exclaimed triumphantly.

Alice said as gravely as she could, "They might go different ways." But she couldn't help thinking to herself, "What dreadful nonsense we *are* talking!"

"She can't do sums a *bit*!" the Queens said together, with great emphasis.

Publications from FairTest

In addition to the *FairTest Examiner*, the National Center for Fair & Open Testing has several other important publications available to meet your needs.

Fallout From The Testing Explosion: How 100 Million Standardized Exams Undermine Equity and Excellence in American Public Schools, by Noe Medina and Monty Neill, is FairTest's study, released in June, 1988 of the extent and effects of standardized testing on children and curricula in grades K-12. (45 pp. $8.95)

Sex Bias in College Admissions Tests: Why Women Lose Out, by Phyllis Rosser, examines the impact of biased standardized university admissions tests on women and their ability to get into college and receive scholarships. (40 pp. $7.95)

Beyond Standardized Tests, Admissions Alternatives that Work, by Amy Allina, takes a look at the successes of several schools that have dropped or deemphasized the use of standardized tests for admissions. (15 pp. $5.50)

The Reign of ETS, by Allan Nairn, the classic 1980 study of the power of the Educational Testing Service. (550 pp $30.00)

None of the Above: Behind the Myth of Scholastic Aptitude, by David Owen, "An angry book (though often a very funny one), and with good reason:the SAT is a scam." (Jonathan Yardley, *Washington Post*) (327 pp. lists at $7.95, $5.95 from FairTest.)

To order these or other FairTest publications, including our informative bi-annual *Updates* use the order form on p. 15. Please note that we prefer pre-payment for amounts less than $25.00.

-Rocky Mountain News Sun., Dec. 4, 1988, Denver, Colo.

EDUCATION NEWS

Fair-testing code fails, some say

Critics decry lack of enforcement

By AMANDA COVARRUBIAS
Rocky Mountain News Education Writer

DICK DAVIS/Rocky Mountain News

Angela Roth, left, and Vereza Hartudo, students at Baker Junior High School in Denver, concentrate during a recent test.

A fair-testing code designed to benefit millions of students nationwide is ineffectual because it lacks provisions for enforcement, say some Colorado testing officials.

The Code of Fair Testing Practices in Education was released two months ago, partly in response to charges that many achievement tests are biased against females, minorities and students at schools where curricula don't reflect the test questions.

But the code itself has generated criticism.

"It may be a good first step, but I don't think it goes far enough," said Gerald W. Bracey, testing director for Cherry Creek Public Schools. "There are no sanctions. Somebody has to have the power to impose sanctions."

The code is particularly significant for Colorado, which has a large number of Hispanic students in the Denver area and throughout the state.

John Fremer, a researcher for the Educational Testing Service and a member of the committee that developed the code, said the point is to inform people, not punish them.

"It's not so much a big stick to beat people up so they will behave properly," he said. "Most of the problems stem from misinformation or just never having learned . . . but I suppose a stick wouldn't hurt either."

The code is sometimes referred to as the Test Takers' Bill of Rights because it aims to answer the question, "What do you have the right to expect from a test?" Fremer said.

It establishes 21 standards in four areas for test developers and users: developing and selecting tests, interpreting scores, striving for fairness, and informing test takers. It also commits test givers to investigate performances of test takers of different races, sexes and ethnic backgrounds "when samples of sufficient size are available."

Colorado testing director Wayne Martin said the Colorado Department of Education already follows many of the guidelines, but the code could be useful to local school districts, he said.

"Hopefully it will do away with some of the potential abuses, which result more from a lack of information rather than outright abuse," he said.

Fremer said the next step is to educate teachers about how achievement tests should be used, because teachers comprise the biggest group of test givers. He acknowledged there is an overreliance on achievement tests, partially triggered by publishers eager to sell their products.

"Testing is more at risk from its advocates than from its critics," he said.

One of the most vocal critics is FairTest, a non-profit organization in Cambridge, Mass., dedicated to fair and open testing. Director Monty Neill is not comforted by the standards.

"While it represents some public indication of a need for reform, in general it's much too vague and lacks enforcement mechanisms," he said.

Neill said he doubts a powerful testing company such as the Educational Testing Service, which publishes the Scholastic Aptitude Test for college admission, would have endorsed the code if it had sanctions.

BEST COPY AVAILABLE

Subscribe to FairTest Publications

FairTest publications will keep you up-to-date on the growing movement for fair & open testing. You will be able to do more than just read about FairTest campaigns —you can help make them successful!

_____ *FairTest Examiner*, published quarterly...Individuals $15/year; Institutional checks $25

_____ *Fallout from the Testing Explosion: How 100 Million Standardized Exams Undermine Equity and Excellence in America's Public Schools*, by Noe Medina and Monty Neill with FairTest staff (45 pages)...$8.95

_____ *Sex Bias in College Admissions Tests: Why Women Lose Out*, by Phyllis Rosser and FairTest staff, 2nd edition (40 pages)...$7.95

_____ *The Reign of ETS*, by Allan Nairn and Ralph Nader (550 pages)...$30

_____ *Beyond Standardized Tests: Admissions Alternatives that Work*, by Amy Allina with FairTest staff (15 pages)...$5.50

_____ *None of the Above: Behind the Myth of Scholastic Aptitude*, by David Owen (327 pp. lists at $7.95)...$5.95.

Updates, issued bi-annually (spring & fall). $7/year/topic:

_____ *Testing Reform Initiatives*
_____ *K-12th Grade Testing*
_____ *University Admissions Tests*
_____ *Teacher Testing*
_____ *Employment Testing*
_____ *Sex Bias in Testing*
_____ *Test Coaching*

Order total _____
Tax-deductible contribution _____
TOTAL ENCLOSED _____
(enclose check to FairTest)

_____ Back issues of *FairTest Examiner* $4 each: Vol___, No.___ (all 7 for $24.00)

_____ **Moving.** I have moved/will be moving. My new address is (Please enclose old label or write old address.):

(Prepayment on orders under $25.00 is greatly appreciated.)

Name ______________________________ Title ______________________________
Organization ______________________________ Phone No. ______________________________
Address ______________________________ Zip ______________________________

Make checks payable to FairTest and mail to: FairTest, 342 Broadway, Cambridge, MA 02139

Southerners to Meet on K-12 Testing Reform

More than two dozen groups and educational leaders from across the South have joined with FairTest to sponsor a Southern Regional Conference on Testing Reform in the Public Schools, to be held March 18 and 19, 1989, in Atlanta, Georgia. This event will bring together civil rights activists, children's advocates and educators to plan campaigns to reduce the misuse of and overreliance on standardized testing.

Among the conference co-sponsors are the Intercultural Development Research Association (IDRA) of Texas, the Atlantic Center for Research in Education (ACRE) of North Carolina, the Mississippi Children's Defense Fund, Kentucky Youth Advocates, the Louisiana Association for Young Children, the Alabama Association for Young Children, the National Coalition of Title I/Chapter I Parents, Robert Brown of Macon, Georgia, Asa Hilliard of Georgia State University, and Constance Kamii of the University of Alabama at Birmingham. Representatives are also likely to come from Virginia, South Carolina, Florida, Arkansas, Tennessee and Oklahoma.

Goals include:

—forming coalitions of testing reform activists, including civil rights, children's rights and education activists, in many Southern states;

—developing a basic strategy for promoting testing reforms locally and statewide;

—creating a regional support network of advocates and experts to assist local and statewide efforts;

—exchanging information on local and state use of standardized tests;

—designing strategies for the systematic collection of information on test use in different states; and

—defining FairTest support for local, statewide and regional campaigns.

This conference springs from the successful testing reform efforts of activists in North Carolina and Arizona (see *FairTest Examiner* Vol.2, Nos. 2 & 3), workshops at FairTest's national conference in June, and widespread grassroots disillusionment with testing. For example, in Georgia, which only this year implemented testing kindergartners for admission to first grade, educators at all levels have already come out in opposition to the exams.

Reform campaigns focusing on such issues as the testing of young, minority, handicapped and special needs students already exist in many states and locales. Bringing these groups together to learn from one another and to plan a stronger and deeper movement is the mission of this conference.

Activists interested in attending the conference should contact Monty Neill or Mike Friedman at FairTest for more information.

FairTest
National Center for Fair & Open Testing
342 Broadway
Cambridge, MA 02139
(617) 864-4810

Address Correction Requested

Bulk Rate
U.S. Postage
PAID
Cambridge, MA
Permit No. 53278

Please subscribe today.
See page 15.

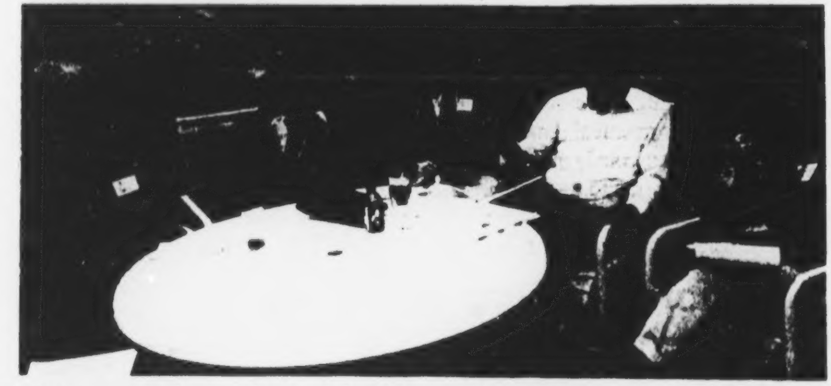
FairTest

EXAMINER

The National Center for Fair & Open Testing Volume 3, Number 2 Spring 1989

NY Decision Leads Way for Testing Reform

"After a careful review of the evidence, this Court concludes that SAT scores capture a student's academic achievement no more than a student's yearbook photograph captures the full range of her experiences in high school."
—Judge John M. Walker



Participants at the Southern Regional Conference on Testing Reform: l. to r. Ann Gould, Arizona; Judy Doebler, Oklahoma; Cathy Grace, Mississippi; Robert Bell, Mississippi; Roz Charlesworth, Louisiana (see story, p. 16).

In a landmark federal court decision, District Judge Walker ruled, in February, that New York's policy of awarding scholarships on the basis of SAT scores alone discriminates against females. Walker also concluded that the New York State Education Department could not justify its discriminatory practice with the argument that any alternative would be difficult to administer. Most states, he explained, run scholarship programs that utilize criteria other than standardized test scores.

Walker's decision has national implications. Since the decision, campaigns and initiatives to limit the use of test scores have proliferated. The National Collegiate Athletic Association (NCAA) has already indicated it is likely to repeal Proposition 42 (see related story, p. 8); legislation has been filed to ensure that the Massachusetts

(continued on page 5)

Georgia Drops Test; Is Texas Next?

The state of Georgia will drop a controversial test that has been used to determine entry into first grade. The kindergarten test, used last spring for the first time, was criticized by teachers, administrators, local school boards and parents as traumatic to children and educationally unnecessary. Because the test was not mandated by state law, the State Board of Education was able to both implement and drop it arbitrarily. The exam will be given this spring for the final time, but cut-off scores will not be used for promotion. The state board plans to implement an alternative assessment.

Georgia joins other states that have recently stopped testing in one or more grades. North Carolina banned testing in grades one and two, Arizona stopped mandated testing in grades one and 12, and Mississippi will drop its kindergarten test.

The state of Texas may soon become part of this trend. With wide backing from education reform groups and the State Board of Education, the Texas Senate has passed legislation to drop its mandated grade one test. The bill is expected to become law. Texas does not require a second grade test, but does test in grade three. Unlike North Carolina, which completely banned testing, the Texas legislation would permit districts to continue using standardized tests.

CONTENTS

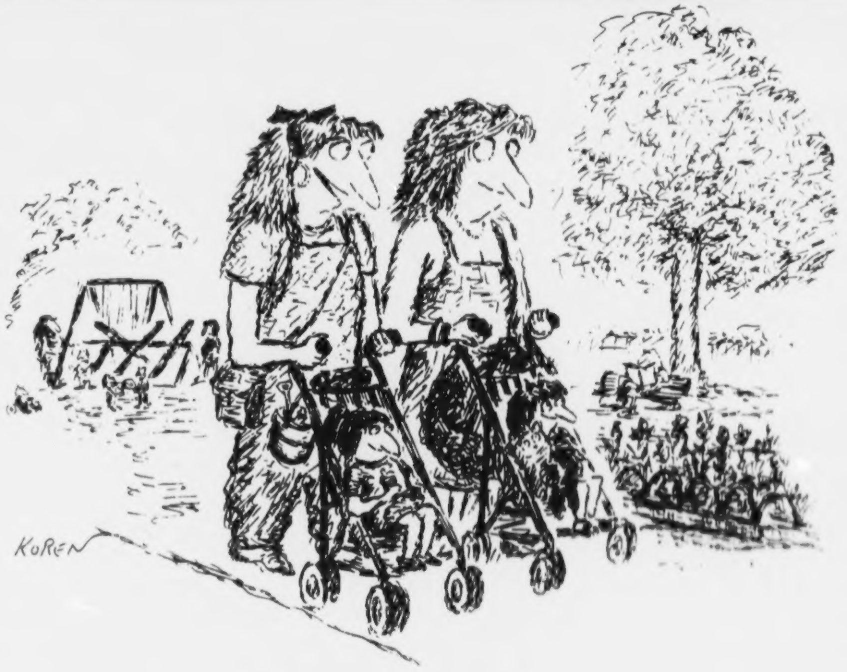
SAT/ACT Optional? 3
TASP Criticized 3
More New Tests 4
Mass. Scholarships 5
Lake Wobegon Revisited 6
NC Testing Alternatives 7
NCAA to Drop Prop. 42 8
SAT Guinea Pigs 9
ETS/CEEB Distort NY Loss 10
R.I. NTE Report 11
Testing and Basal Readers 12
Southern Conference 16

TM017189

Litigating Employment Testing Cases

An important new resource is now available for lawyers and other advocates working for fair employment practices. "Why Plaintiffs' Counsel Challenge Tests, and How They Can Successfully Challenge the Theory of 'Validity Generalization'" is a 30-page overview written by Richard Seymour, Director of the Employment Discrimination Project of the Lawyers Committee for Civil Rights Under Law. The paper covers a variety of approaches to tackling test-based discrimination and suggests practical remedies.

The author has generously made reprints available for distribution to interested activists. To obtain a copy of Seymour's paper, send $.65 in stamps to FairTest.



"How are her scores?"

National Commission Moves to BC

With the departure of commission chair Dr. Bernard Gifford, the National Commission on Testing and Public Policy has moved its offices from the University of California at Berkeley to Boston College's Center for the Study of Testing, Evaluation and Education Policy (CSTEEP). George Madaus, Director of CSTEEP, is the new chair of the Commission.

You can reach the Commission by writing National Commission on Testing and Public Policy, Boston College CSTEEP, Chestnut Hill, MA 02167.

The National Center for Fair & Open Testing (FairTest)
342 Broadway, Cambridge, MA 02139; telephone (617) 864-4810.
April, 1989

Staff: Mike Friedman, Publications Manager; Monty Neill, Managing Director; David Riker, Office Assistant; Bob Schaeffer, Communications; Cinthia Schuman, Executive Director; Sarah Stockwell, Admissions Tests Coordinator.

Board of Directors: Prof. Denise Carty-Bennia, Northeastern University Law School, Chair; Dr. Pamela George, NC Central University; Paul Pottinger, New Manhattan Corp.; William Robinson, Dean, D.C. School of Law; Bob Schaeffer, Treasurer; Steve Solomon, Esq.; Chuck Stone, *Philadelphia Daily News*; Ron Vera, Esq., Barbosa & Vera; John Weiss, President.

Partial funding for FairTest's work comes from the Ford Foundation and the Rockefeller Family Fund.

Quote of the Month

"The short right answer has come to dominate education, but only because of the mania for teaching and testing by the most mechanical means possible—'objectively.' The theory is that to teach and evaluate learning in any way except by numbers would be unscientific and unreliable, and to allow scores and grades to be allocated on the basis of personal judgment would be 'subjective' (a pejorative term) and biased. Ironically, the endeavor to remove the personal, and possibly prejudicial, from educational evaluation has resulted in the totally arbitrary and distorted procedure of teaching and testing only those things that can be cleanly scored right and wrong, and counted. The cost of removing human error has been the removal of humanity and the reduction of education to trivia."

—Frank Smith, reported in *Rethinking Schools*, December/January 1987/1988.

Correction:

In the last issue of the *Examiner*, the cartoon on page 3 should have been credited to the National Student News Service. We apologize for this omission.

University Admissions Testing

California Bill Would Make SAT and ACT Optional in Admissions

Legislation recently introduced in the California state senate would make SAT and ACT scores optional for admissions to schools in the University of California (UC) and California State University (CSU) systems. This first-of-its-kind legislation represents an important step toward halting the overuse and misuse of these tests. The bill, Senate 1599, was filed by state Senator Art Torres in late March. Public legislative hearings on the two-year bill will be conducted this summer.

The Reverend Jesse Jackson expressed his support for the bill at a rally of 10,000 students for educational rights on April 10. Jaime Escalante, the subject of the movie *Stand and Deliver*, who teaches Advanced Placement Calculus at Garfield High School in Los Angeles also has endorsed the bill. In addition, the Southern and Northern California Civil Rights Coalitions, which include groups such as the NAACP, the American Civil Liberties Union the Mexican American Legal Defense and Education Fund (, the Latino Issues Forum and the National Organization for Women, are likely to support S.1599.

Test manufacturers are very concerned about the proposal and are stepping up their argument that the tests are not a problem. According to the College Board, for instance, legislators and groups that support the bill would better serve disadvantaged students by encouraging them to increase their academic courseload and prepare for the test. However, the sponsors of the California legislation recognize that substantial evidence (much of it by the test companies' own researchers) demonstrates that the SAT and ACT do not merely *reflect* differences in academic preparation, but actually *exaggerate* these differences.

S.1599 also asks the UC and CSU Academic Senates to study graduate admissions tests, including the LSAT, GRE, GMAT, MCAT and the Dental Admission Testing Program Test. The Academic Senates would then report their findings to the California Post-Secondary Education Commission (CPEC) for recommendations.

"My problem is I did great on the SAT, but classes in college are nothing like that."

Texas Criticized on TASP

The Texas Higher Education Coordinating Board has recently instituted a basic skills test for college students with cut-off scores that will likely fail more than three-quarters of Blacks and two-thirds of Hispanics. Students who fall short will be barred from taking upper level courses, thus suspending their education, until they pass all three sections of the test. Critics such as the Mexican American Legal Defense and Education Fund (MALDEF) vigorously oppose the high cut score and fear that many low scorers will not be able to complete their education.

The Texas Academic Skills Program (TASP) test consists of three multiple choice sections, math, reading and writing (which includes an essay). Students who fail the test cut-off will be forced to enroll in as yet unfunded remedial courses. The test, produced by National Evaluation Systems of Amherst, Massachusetts, will cost students $24.00 each time they take it.

Though the Board expects very high failure rates, particularly for Hispanic and Black test-takers, Albert Kauffman of MALDEF argues they will be even higher than official estimates. Based on the answers to sample questions given last fall, Kauffman says that as many as 93% of Blacks, 87% of Hispanics and 74% of Whites will fail the TASP, a 77% overall failure rate. He urged the Board to lower the passing score and gradually phase in the test over a four-to-six year period. "Making the test an absolute bar to progress is what really concerns us," Kauffman said.

The first official TASP test was administered in March to a small group of students, mostly teacher-training candidates. Results were better than MALDEF expected, but Kauffman said that the sample was far from normal. He still has reservations about the TASP: "We have some concerns whether the test reflects the basic skills of Blacks and Hispanics as

(continued on page 13)

University Admissions Testing

Learning Disabled Students and the SAT

Editor's note:

Like members of minority groups, women and students from lower-income families, students who require a "non-standard" administration of the SAT face unfair hurdles to higher education.

"Special Administration" SATs, those that are untimed, large type, audio cassette, in Braille, etc, are flagged as such on score reports to colleges. These are not considered "standard" and are therefore not accepted on par with other scores by many colleges and universities.

Beth Lohman, an educational consultant from Wayne, Pennsylvania, wrote to tell us of another unfair hurdle faced by Learning Disabled students.

Whenever Learning Disabled (L.D.) students apply to take the SAT or other ETS college admissions test with extra time, they are required to present either recent test results or a current individualized education plan (IEP) to certify their learning disability. These regulations, which seem fair on the surface, often restrict some of the very students who have successfully dealt with their L.D.

It is an accepted fact among those involved with L.D. students that there is no "cure" for a learning disability. Successful remediation means that the student has learned compensatory skills well enough to require no further remedial help. A student, for example, may have developed good comprehension skills but still reads very slowly. He or she may have completed advanced math courses but had to use mnemonic devices to learn and utilize theorems. An L.D. student often requires additional time, but many can then perform very well.

A student who has had successful remediation typically does not have what ETS defines as "current" documentation of their L.D. — precisely because they no longer need remediation. The ETS requirement forces a student's parents to either pay for outside testing of their child (an expensive proposition at $200-$700), or talk the school into providing psychological testing.

Schools are often reluctant to do this because they lack personnel and funds and because it's not information that they need in order to provide services to the student. Counselors then can feel pressured to circumvent the regulations either by not having the required documentation in the student's file when certifying the student as L.D., or by getting out-of-date documentation.

ETS should educate itself as to what being L.D. really means for a lifetime of learning. As it stands, ETS' documentation requirements are just one more unnecessary hindrance to the L.D. student trying to pursue a college degree.

—Beth Lohman

PEANUTS By Charles M. Schulz

IT'S GOING TO BE A HISTORY TEST, AND I NEED YOUR HELP
3-6
FORTUNATELY, I THINK IT'LL PROBABLY BE MYSTICAL CHOICE
MULTIPLE CHOICE
WHATEVER
Schulz

Coming Soon to A Classroom Near You

In the wake of criticism by testing reformers and increased scrutiny of admissions tests, several large testing concerns have announced revisions or major studies of their tests. However, none of the changes appear to address the fundamental biases and other flaws in the tests, making the changes largely cosmetic.

Last fall, the Educational Testing Service announced that by 1992 the National Teacher Exam (NTE) will be replaced (see FairTest *Examiner*, Vol. 2, #4) and that a major review of the SAT will be conducted. The American College Testing Program (ACT) has announced substantial changes in its admissions test, with the new test set for its first administration this Fall (See FairTest *Examiner*, Vol 3, #1).

Other tests that have announced revisions or reviews recently are the Medical College Admissions Test (MCAT), which will significantly change its format, and the Law School Admissions Test (LSAT), a new version of which will be introduced in June, 1989. The Law School Admissions Council (LSAC) also plans to conduct a major review to determine whether the LSAT is gender biased. The last such review yielded "inconclusive" results, according to an official of the LSAC.

The new MCAT, to be introduced in 1991, will be reduced from six sections to four plus a writing sample, resulting in 80 minutes being shaved off of the 9 1/2 hour test. In the face of declining medical school applications, the changes are evidently intended to encourage students from broader educational backgrounds to consider entering medical school. The current format of the test has been criticized for relying too heavily on rote memorization and not enough on "critical thinking" and "problem solving skills." Evidence that the proposed changes actually measure these skills any better than did the old MCAT has not been presented.

University Admissions Testing

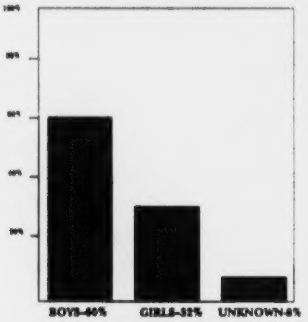
Is Massachusetts Next?

FairTest Targets Another SAT Misuse

Following on the heels of the New York court victory, legislation has been filed in Massachusetts to ban the use of SAT scores in the state's Senatorial Honors Scholarship program. Massachusetts is now the only state in the nation which continues to make scholarship decisions solely on the basis of SAT scores. As a result, only 32% of Senatorial Honors Scholarships went to girls last year, while boys received at least 60% of the awards. The gender of the remaining 8% could not be determined from official lists. "This practice violates federal anti-discrimination laws and directly conflicts with the New York court decision," said Cinthia Schuman, FairTest's Executive Director, testifying in April before the Joint Committee on Education, Arts and Humanities of the Massachusetts legislature. FairTest spoke in support of an omnibus bill that would create a more equitable scholarship program.

To make sure *this* year's class of high school seniors don't suffer discrimination, Senator Sal Albano and Representative Sue Tucker have filed an additional bill, "An Act to Ensure Fairness in the Senatorial Honors Scholarship Program." This legislation would require the state to use a broad range of factors in determining 1989 scholarship awards, including academic, leadership and extracurricular achievement, but not including SAT scores.

Massachusetts Senatorial Honors Scholarships 1988-1989



New York Victory *(from page 1)*

Senatorial Honors Scholarships will no longer be awarded using only SAT scores (see story, above); and lawyers are looking into litigation against the National Merit Scholarship Corporation, which relies solely on PSAT scores to select semifinalists for National Merit Scholarships. In addition, legislation was recently filed in California to make test scores optional for admission to the state's university systems (see story, p. 3)

In New York, relying solely on the SAT resulted in over 70% of top scholarships going to boys. In response, the American Civil Liberties Union Women's Rights Project and the New York Civil Liberties Union filed suit on behalf of ten female high school seniors, the National Organization for Women (NOW) and the Girls Clubs of America. The plaintiffs argued that use of the SAT discriminated against girls for two reasons: 1) the SAT does not measure high school performance and, even if it did, 2) the SAT underpredicts females' academic ability as compared with males'.

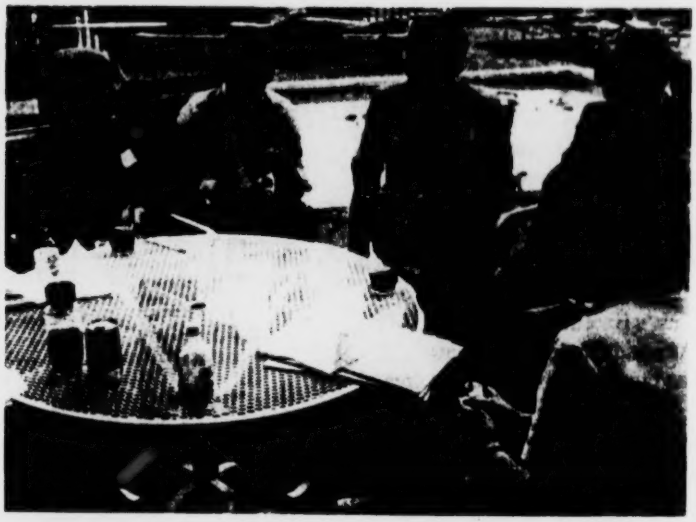
The issue of whether the SAT measures past performance was critical to the case because the New York scholarships are supposed to reward students who have demonstrated academic achievement in high school. The state claimed the SAT reflects the high school curriculum.

However, Judge Walker found that SAT scores "do not mirror high school Math and English classes" and fail "to provide *any* measure of what a student learns in foreign language, science, and social studies courses. Moreover, there can be no serious claim that a test given on one single morning can take into account a student's diligence, creativity and social development and work habits in that student's environment — all part of high school achievement."

Walker also concluded that the SAT underpredicts for females: "the probability, *absent discriminatory causes*, that women would consistently score 60 points less on the SAT than men is nearly zero" (emphasis added). After isolating the effects of gender from other variables such as ethnicity, socio-educational status (parental education), high school classes, and proposed college major, he found that, "under the most conservative studies presented in evidence, even after removing the effect of these factors, at least a 30 point combined differential remains unexplained."

The court decision is a great victory for FairTest, which has long contended that the SAT is biased against women as well as minorities. FairTest's 1987 report, *Sex Bias in College Admissions Tests: Why Women Lose Out*, by Phyllis Rosser, initiated work on gender bias in admissions tests and the National Merit Scholarships. FairTest also worked with the New York Public Interest Research Group in analyzing the New York scholarship awards in 1987, which led to the suit.

Grade School Testing



Participants at the Southern Regional Conference: (l. to r.) Mary Jo Pollman, Alabama; Blair Horner, New York; Lewis King, Alabama; Bob Witherspoon, Washington, DC; Constance Kamii, Alabama.

Lake Wobegon Revisited

A new national study by the Center for Research on Evaluation, Standards and Student Testing (CRESST) at the University of California at Los Angeles has concluded that the "Lake Wobegon effect" does exist: all states and most students *are* above average on nationally normed standardized tests. It found that the major causes for this strange occurrence are "teaching to the test" and outdated norms.

In a 1987 study done by Friends for Education, West Virginia physician John Jacob Cannell first reported what became termed the "Lake Wobegon effect," after Garrison Keillor's fictional town where "all the children are above average." Cannell's report created controversy and was attacked by test companies for faulty statistical procedures.

According to Robert Linn, a co-author of the new study, for all 35 states studied, "the overall percent of students above the national median is greater than 50 in all the elementary grades in both reading and mathematics for each of the three years studied," essentially confirming Cannell's report. Linn added, however, that "the findings are not as extreme as Cannell suggested."

Lorrie Shepard, another co-author, added, "Intense pressure on educators to improve scores sets the stage and increases the incentives for the various types of teaching to the test." She noted that teachers can teach items they remember from previous test administrations. Knowledge of only a few items can lead to substantial percentile increases.

"Instead of thinking about the impact of Cannell's analyses as the 'emperor's new clothes' metaphor, I propose we think about deposing the emperor," concluded Eva Baker, director of CRESST. Baker called for using alternative, improved assessments. (Quotations from *Education Week*, 4/5/89.)

Appropriate Literacy Evaluation

On May 13, 1989 in New York City, some 100 people are expected to participate in a conference, co-sponsored by FairTest, to exchange ideas on emerging techniques in alternative literacy evaluation. The conference will highlight "appropriate assessments" which could replace standardized testing. The other co-sponsors include Advocates for Children of New York, American Reading Council, Educational Priorities Panel, New York Public Interest Research Group (NYPIRG), Public Education Association, and the Writing Center of Lang College at the New School for Social Research.

Following a welcome by Representative Major Owens (D-NY), Monty Neill of FairTest and George Madaus of Boston College will deliver a critique of standardized testing.

In the morning, people working on alternatives will present their ideas in small group sessions. In the afternoon, workshops will discuss and critique the various models. The closing session will feature Ted Chittenden of ETS, Susan Lytle from the Literacy Research Center at the University of Pennsylvania, and Deborah Meier of Central Park East Secondary School in New York City. This session will suggest ways to further the understanding, availability and use of appropriate literacy evaluation. Conference sponsors hope that a handbook of alternatives can be produced and widely disseminated.

The conference will be held at the New School for Social Research, 65 West 11th St., 3rd Flr., New York, from 9:00 am to 5:00 pm, with wine and cheese to follow. For further information or to register, call Joan Harris at NYPIRG, (212) 349-6460.

Grade School Testing

Opposition to School Testing Grows

The **National Association of Early Childhood Teacher Educators** (NAECTE) has voted to go on the offensive against the misuse of standardized tests. In a November resolution, NAECTE reported that "a number of highly questionable practices have resulted from the widespread use of screening, readiness and achievement tests" on young children. Among the negative results of such test misuses are denial of entrance into regular programs, group segregation, increased retention, and excessive use of drill on isolated skills by teachers pressured to raise scores.

According to NAECTE, these practices are not "consonant with the body of knowledge that explains how children learn," and stifle young children's motivation to continue learning. The group therefore resolved "to compile and disseminate appropriate resources...to individuals responsible for making decisions about classroom practice." For more information contact Dr. Harlan S. Hansen, University of Minnesota, 125 Peik Hall, Minneapolis, MN 55455.

The **National Association of School Psychologists** has issued a position statement on "Student Grade Retention" which reports that "retention decisions cannot be validated using any standardized or competency-based tests." The statement notes that "retention can negatively affect achievement and social-emotional adjustment" and commits NASP to finding methods to help practitioners "avoid the inappropriate use of retention" and encourage "interventions other than retention" (NASP, 808 17th St. NW, #200, Washington, DC 20006).

North Carolina Adopts Alternatives

The North Carolina Board of Education has approved alternative assessments for the first and second grades. The new procedures will replace the standardized tests dropped in 1987.

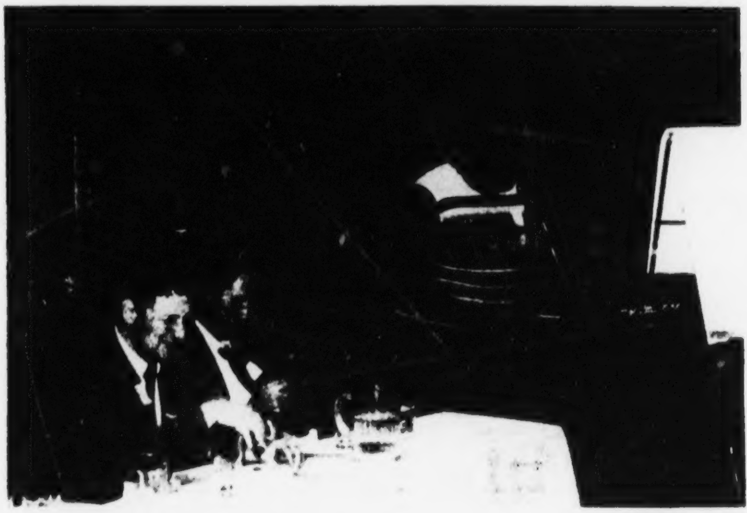
The new assessments are based on teacher observations of student classroom performance in communications and math and will include samples of the children's work. These "appropriate assessments" are designed to provide a rich source of information to teachers based on the children's actual work and progress in the classroom. New forms and procedures have been designed to guide teachers in assessment.

According to Norma Kimzey, who helped develop the new model, the next step needed is to report this information to parents in a useful form. Later, education officials will have to develop the means to collect data on student achievement across districts and the state. She is concerned, however, that in so doing they may end up breaking knowledge into little bits and pieces, thereby reproducing one of the flaws of standardized tests.

Recommended Reading

The May 1989 issue of *Phi Delta Kappan* will feature an article by FairTest's Monty Neill and Noe Medina on the problems associated with the overuse and misuse of standardized tests. The article is part of a larger section discussing alternatives to standardized testing that will include an analysis of the criteria alternatives ought to meet and examples of various new methods of assessment that can replace testing. (Copies are $2.50 ppd. from PDK, P.O. Box 789, Bloomington, IN 47402-0789.)

Sue Bredekamp of the National Association for the Education of Young Children and Lorrie Shepard of the University of Colorado have co-authored "How Best to Protect Children from Inappropriate School Expectations, Practices and Policies," in *Young Children*, March 1989. It includes an excellent critique of the misuse of standardized testing on young children and a helpful bibliography. The issue is $4, the article $2, from NAEYC, 1834 Conn. Ave. NW, Washington, DC 20009.



L. to r. Bob Schaeffer and Monty Neill of FairTest, and William Tate, Jr., Detroit, at the Southern Regional Conference.

University Admissions Testing

NCAA Moves to Drop Prop. 42; Test Misuse Frames Debate

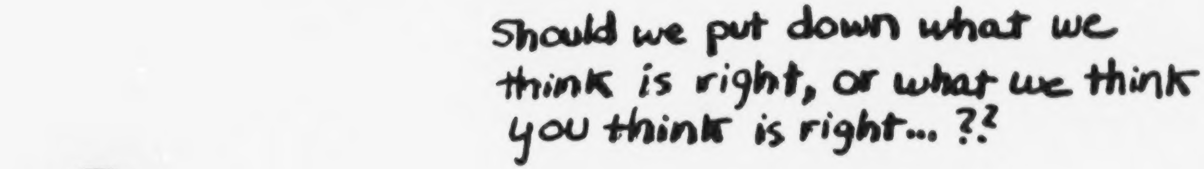
In response to pressure from many coaches, college presidents, and FairTest, the National Collegiate Athletic Association (NCAA) has begun backing away from Proposition 42. The rule, adopted at the NCAA's January national meeting, would deny athletic scholarships to students who fail to obtain both a minimum high school grade point average and a score of 700 on the SAT or 15 on the ACT. The majority of those who would be denied scholarships would be Blacks and Hispanics who score below the test score cut-off.

Reacting to the NCAA's adoption of Prop. 42, FairTest called on the NCAA to either justify its use of test scores or repeal the plan. In response, Dick Schultz, Executive Director of the NCAA, wrote that the NCAA Presidents Commission and the NCAA Council "have declared their intention to oppose Proposition 42 when it comes before the January 1990 NCAA Convention. Since the proposal has an effective date of August 1990, its repeal would mean that no student-athletes had been affected." Under the current NCAA rule, Proposition 48, students who do not reach the cut score can still receive scholarships, but may not play during their first year.

The key issue of Prop. 42 is whether students who would be barred from receiving scholarships — and therefore, in many cases, college educations — on the basis of test scores, are incapable of doing college work. A study by Timothy Walters at the University of Michigan at Ann Arbor has documented that the vast majority of those who would be denied scholarships do, in fact, successfully complete their first-year coursework and go on to graduate (see *Examiner*, Vol. 3, # 1).

While the NCAA is acting responsibly to reevaluate its position, test-makers have done little to block the misuses of test scores, which are contrary to their guidelines, besides offering to assist the NCAA in coming up with new regulations.

Richard Ferguson, President of the American College Testing Program (ACT), never answered FairTest's letter asking that his firm withhold scores from colleges and universities which would use cut scores to exclude students in connection with Prop. 42. Gregory Anrig, President of the Educational Testing Service (ETS), did criticize both Prop. 42 and Prop. 48 under prodding from FairTest Board Chair Denise Carty-Bennia in a "CBS Morning News" TV debate. But he later protested that ETS cannot withhold test scores since it is under contract with the College Board to provide these services. FairTest maintains that ETS should prevent the misuse of its product, including, if necessary, amending its contract with the College Board.


Should we put down what we think is right, or what we think you think is right... ?!
APRESS

University Admissions Testing

SAT-Verbal Measures Test Wiseness

Reading comprehension passages on the SAT-Verbal section measure test-taking abilities more than the reading skills they claim to assess, according to a new study by researchers at the University of Georgia. The experiment was inspired by David Owen, author of *None of the Above*, who claims that reading the passage is often unnecessary for answering reading comprehension questions correctly. Professor Stuart Katz, who led the experiment, found that students were able to answer nearly 40% of reading comprehension questions correctly without reading the passages themselves. Moreover, students' performance on this task correlated closely with their scores on the SAT-Verbal section, suggesting that SAT-Verbal scores are heavily influenced by test-wiseness as opposed to any real verbal ability.

In an *Atlanta Journal* story, Nancy Burton of ETS said "reading comprehension is one of those skills that develops over time...the reading passages simply rely on your ability to comprehend and analyze what you read." She doubted that a student's test-taking skills or knowledge of the reading topic would result in their answering even 20% (the rate of chance) of the questions correctly without the passage.

"Theoretically, if you eliminate the passage, you should not be able to perform better than chance," Professor Katz told the *Journal*. "But people are doing much better than chance. They're test wise — they're able to use tricks of various sorts to find the correct answer. This can't be due to reading comprehension skills because the passage is not there." For a copy of the study, contact Dr. Katz at (404) 542-3039.

SAT Guinea Pigs

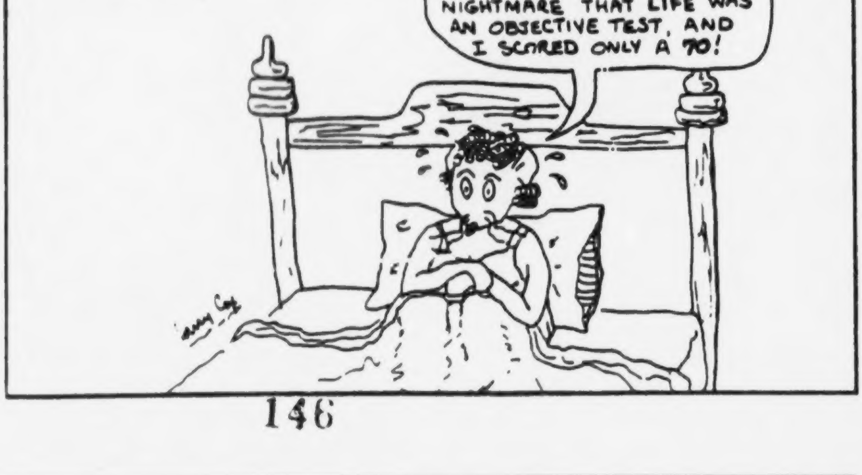
In response to growing criticism of the SAT, the Educational Testing Service (ETS) has begun testing both new types of questions and alternatives to the current experimental section. Some ETS experiments are using students as study subjects during the school day and others are administered before or after regular SAT exams.

One study administered new types of SAT items, including writing samples and open-ended questions, to high school juniors during their regularly scheduled class time. Another study, carried out in Washington, Illinois, Ohio, Michigan, Connecticut and New York as a part of the March 11 test administration, asked students to arrive a half hour early to take an extra section of SAT questions. The College Board claimed the experience "will give you a unique chance to 'warm up' before taking the SAT."

In a letter to ETS Study Director Cathy Wendler, Louis Bougie, Secretary of the South Central Area Guidance Organization and test center supervisor of Branford High School in Connecticut, protested that "(t)o ask students to participate in a 'special study' to develop questions for further tests strikes us as placing a burden on students on an already stressful day. To suggest to students that the volunteers have been given 'a unique chance to "warm up"....' seems to raise questions of reliability, validity and integrity in the entire test....(C)ould 'warming up' backfire?"

According to Wendler, this study of pre-testing is being undertaken in response to concerns from students, parents, educators and legislators about the SAT's *current* experimental section.

Legislation under consideration in New York would establish a "Test-Takers Bill of Rights" including provisions to protect students from being forced to pay for the privilege of serving as ETS' guinea pigs. Last year ETS testified at a legislative hearing in Albany that it was not feasible to develop questions except by including an experimental section in every SAT exam. Now they are doing exactly what they said they could not. But the new system shows no more respect for students, the test's real consumers, than the old system does.


Of Mice and...
WHEW! I'M LETTING TEACHING GET TO ME. I HAD A NIGHTMARE THAT LIFE WAS AN OBJECTIVE TEST, AND I SCORED ONLY A 70!

University Admissions Testing

ETS and College Board Distort SAT "Sex Bias Scholarship" Case

(FairTest has submitted the following article to other publications.)

In the wake of the landmark New York decision, the Educational Testing Service (ETS) and the College Board (CEEB) have launched a massive public relations campaign to minimize the damage to their lucrative testing business. In their scramble to cast the decision in the best possible light, ETS and the CEEB have resorted to a series of deliberately misleading statements, half truths and outright falsehoods.

For example, both firms now claim that they had always argued against using the SAT as the sole factor to award scholarships. A College Board letter sent to educational leaders the week after Judge Walker's ruling said, "ETS and the College Board have always recommended that the SAT be used along with other criteria (where available) in college admissions decisions and scholarship selections . . . "

That statement is simply not true. In a letter sent just last September to Blair Horner, Legislative Director of the New York Public Interest Research Group (NYPIRG), College Board President Donald M. Stewart wrote that the SAT used alone was the "best" method for awarding scholarships in New York. College Board spokesperson Fred Merino echoed this position in an article in the February *APA Monitor*. Perhaps the retroactive change in the College Board's position comes from a recognition that the earlier statements of both Dr. Stewart and Mr. Merino were in direct conflict with the College Board's own *Test Use Guidelines*.

ETS is also distorting the facts to its own advantage. Soon after Judge Walker's decision, Warren Day of ETS told the *Wall Street Journal* that ETS had backed the plaintiffs in a brief it filed in the New York case. But all the arguments in the joint ETS/College Board *amicus curiae* (friend of the court) submission supported New York State's sole use of the test.

Specifically, the ETS/College Board brief claimed that use of the SAT for "scholarship selection purposes" was valid. It is precisely this contention that Judge Walker's decision contradicts. The claim that the court supported the ETS/College Board position is a triumph of public relations puffery over accuracy.

The protest by ETS and the College Board that "the judge is clearly not taking any kind of anti-SAT stand" is also untrue. Judge Walker's decision is clearly critical of the SAT. His opinion states "the probability, *absent discriminatory causes*, that women would consistently score 60 points less on the SAT than men is nearly zero" [emphasis added].

ETS and the College Board should stop wasting their time (and test-takers' fees) on a campaign to obscure Judge Walker's message. Instead they should act responsibly by overhauling the SAT to eliminate its pervasive biases and move effectively to block misuses of test scores. The first step toward genuine reform is to begin telling the public, policy-makers and the media the whole truth about the SAT's flaws.

Coaching Tip for TOEFL

As fast as testing expands, so do test-preparation services. Here's a tip to help beat TOEFL, the Test of English as a Foreign Language, an ETS exam given to thousands of foreign students each year.

TOEFL test-takers are supposed to listen to an oral statement, then read a printed question and answer it. Test-coachers, however, suggest that the student ignore the entire first item, and read the *second* question and answers while the first passage is being read. They will then be prepared to listen to the second passage knowing in advance what the question will be. Since they can now answer it more quickly, the remaining answer time for question 2 can be used for reading the question and answers to the third item, and so on.

By following this strategy, test-takers can avoid being caught by surprise. This is basically a variation on a well-known test-taking technique: reading the question and answers before reading the passage. Of course, mastery of this skill means nothing about the ability to read or write English. But it can increase the standardized test score. In many cases, that's all that matters.

Male/Female Test Differences Narrow—Except on SAT

Test score gaps between men and women in math, verbal, science and spatial skills are declining, according to recent studies by Marcia Linn, University of California, Berkeley, and Janet Hyde, University of Wisconsin-Madison. Linn and Hyde point out that the SAT is an exception to the trend. The SAT-Math section exhibits much larger differences between men and women than any other similar test, and, Linn contends, its test questions are biased in favor of white middle-class males.

Linn and Hyde noted in the March *APA Monitor*, "It is possible to eliminate or exaggerate gender differences by selecting test questions with contexts favoring males or females." This raises fundamental questions about choices made in test construction. Women do better on math problems concerning aesthetics, interpersonal relationships and traditionally female tasks, such as sewing, and men do better at questions involving measurement, sports and science. For copies of the papers, call Marcia Linn at (415) 643-6379.

The Providence Journal

PROVIDENCE JOURNAL-BULLETIN: SINCE 1829 CITY FINAL THURSDAY, FEBRUARY 23, 1989/35 CENTS/

Panel urges moratorium on teacher competency test

By ANN MACARI HEALEY
Journal-Bulletin Education Writer

PROVIDENCE — A legislative commission has urged Rhode Island to stop requiring prospective teachers to pass a national teaching exam until the state can prove a correlation between failing the test and incompetence in the classroom.

The commission's proposal is to continue giving the National Teachers Examination for five years and to allow those who fail it to teach with provisional certificates. Teachers who get such certificates would have their performance monitored to see if poor test results foreshadowed an inability to teach effectively.

Passing the standardized test has been a requirement for new teachers in Rhode Island since 1986, when the Board of Regents for Elementary and Secondary Education joined a growing national movement to improve teacher quality and adopted the exam.

But after protests from minority organizations and the Rhode Island affiliate of the American Civil Liberties Union, the General Assembly in 1987 appointed an eight-member commission to determine whether the test is biased against blacks and Hispanics.

The commission's answer: Yes.

"Effectively, the test systematically discriminates with respect to racial ethnic groups commonly accepted as being underrepresented in the profession," according to the commission.

But the commission did not limit its recommendation to minorities.

"If the test is not valid for minorities, chances are it's not a valid tool" for assessing anyone's teaching skills, said commission member Frank Walker, director of personnel and civil rights at the state Department of Education.

The legislative commission, which includes several educators, suggested that it or the Education Department should devise a way of assessing a teacher's classroom performance in relation to test results.

If the evaluations show that poor testing means poor teaching, then the exam should be used "as a screening device" for certifying new teachers, said the commission. And people who fail the test would not be certified to teach in Rhode Island.

But if no correlation is apparent, the commission said, the test should be discontinued.

"We really do want to create an environment where we can have competent minority teachers in the classroom," said commission chairman Levi Adams, Brown University's associate vice president for biology and medicine.

Recommendation applauded

The ACLU and FairTest, a Cambridge, Mass., organization that seeks to identify culturally biased testing throughout the country, applauded the commission's stance.

"It's a good move," said FairTest managing director Monty Neill. In previous studies, he said, "no one has found any correlation between test scores and performance evaluation."

Twenty-four states use the National Teachers Examination, which tests communication skills and general and professional knowledge. Its author, the Educational Testing Service in New Jersey, recently announced that it will substantially revise the exam.

In 1986 and 1987, about 800 aspiring teachers took the exam in Rhode Island. The test's passing score, established by J. Troy Earhart, commissioner of elementary and secondary education, is among the country's highest. About 75 percent passed the test. Test results for 1988 are not available.

Each part of the test can be taken separately. In Rhode Island, 23 percent of whites and 46 percent of minorities failed communication skills. In general knowledge, 16 percent of whites and 41 percent of minorities failed. In professional knowledge, 9 percent of whites and 20 percent of minorities failed.

The commission said that so few minorities — about 28 — took the test in Rhode Island that the comparisons may be misleading. But it said that the results "are consistent with the national experience."

Earhart declined to comment on the report until he reviews the recommendations.

Regents chairman Augustine Capotosto Jr. said the board, concerned about the test's validity, asked the Education Department last year to study adding performance evaluation to the certification criteria. That report, he said, should be completed about the middle of this year.

"The board recognized that the (National Teachers Examination) is not an end-all but just another piece of evidence in a spectrum of evidence to determine quality of teachers for certification purposes," Capotosto said.

Increasing concern

The poor showing of minorities on the exam has become increasingly disturbing, educators say, because the pool of minority teachers continues to shrink as the numbers of minority students grow. They are projected to account for one-third of the public school population by 2000.

In Rhode Island, 14 percent of the state's 133,585 students this year are minorities. In 1986, according to statistics compiled by the Journal-Bulletin, the state had 120 minority teachers, about 2 percent of the teaching force.

"There is a great deal of anxiety within the minority community with this particular issue," and some will not even take the test, said commission member Juan Francisco, a Hispanic community leader and a member of the Board of Regents.

In its report, the commission also expressed concern that the test could disproportionately exclude skilled artists, musicians and craftsmen without traditional teacher-education backgrounds.

The commission will issue its final report in September 1990. It is hoping, Adams said, for a response from the Department of Education by the end of the month.

BEST COPY AVAILABLE

Testing and Basal Readers

Report Card on Basal Readers, by Kenneth S. Goodman, *et al.* (Katonah, NY: Richard C. Owen, 1988, 167pp., $9.90 postpd.)

Basal readers are the primary tool for teaching reading to most children in the U.S. As awareness has grown about the problems of illiteracy and limited literacy, more attention is being directed to these tools.

This report, prepared for the National Council of Teachers of English, joins *A Conspiracy of Good Intentions: America's Textbook Fiasco* in arguing that it is time to reduce our reliance on basals. Both books are valuable contributions to understanding the reading crisis in the US, the role of standardized testing in the crisis, and possible ways out.

Basals, argue Goodman, *et al.*, strip language of pleasure and often of meaning, assume that learning to read is the sum of a set of discrete components that can be separately taught and tested, and presume teachers are merely technicians.

The basals also fit hand-in-glove with standardized testing. Typically, each chapter or section concludes with multiple choice questions. The children's ability to answer these demonstrates whether they have mastered the content contained in the section. Basal questions are akin to the questions on standardized tests, though they are not subject to even the minimal reliability and validity studies performed on the standardized exams. The *Report Card* finds, "There is no evidence, in their design or in their apparent development, that they meet theoretical criteria for the type of testing they represent" (p. 114).

The assumptions about learning carry over from basal readers to standardized tests and back to basals in a reinforcing circle. The results (limited literacy, dislike of reading, and school failures and drop outs) all demonstrate that the twin engines of reading instruction - basals and standardized tests - need replacement.

Excerpts from *Report Card on Basal Readers*

On Oversimplification:

"The Economy Level B test attempts to be flexible in testing colors by using the phrasing 'decide which picture shows something which is often (color 'x').' This introduces questions of how much 'often' might be. For something that is often green, the choices are [black and white] line drawings of a carrot with a leafy top, a hand, and a leaf... [A] child could hypothesize that the carrot top is more often green than leaves are, since carrot tops are always green in the supermarket but leaves change color with the seasons" (p. 116).

On Alternate event sequence:

"When pupils are asked to arrange a sequence of events in order there is often a variety of orders possible. For instance...:

Dad will cook the eggs in a pot.
Dad will eat the eggs.
Dad will shop for eggs.

Dad will shop for eggs.
Dad will cook the eggs in a pot.
Dad will eat the eggs.

"Both sequences are possible...This example also illustrates another pervasive problem. The language of items is not equally familiar in all American dialects ... This rather urban, middle-class terminology [e.g., 'shop for'] would penalize other learners" (p. 119).

On Lack of context in items testing context (p. 120):

"Threads were hanging from the *tattered* edge of the beach towel. ( ) torn ( ) neat ( ) repaired"

[Harcourt Brace Jovanovich (HBJ), Level 11,"Context Clues"]

On Lack of clarity of the cause/effect relationship (p. 120):

Is this cause or effect?

"4. I'm taking care of the corn plants because *corn is my favorite vegetable.*
cause effect"

[HBJ, Level 11, "Context Clues"]

"These items seem problematic as exemplars of cause and effect relationships."

On Stilted Language:

Finally, the basal tests do not test "the reader's ability to comprehend natural language texts. Rather, what is being tested is the ability of pupils to successfully respond to 'skill items,' including comprehension 'skills.'" The same can be said for standardized "reading" tests. The language in both types of exams is markedly reduced and synthetic even by comparison with the limited and stilted language in the basals themselves. An example from Ginn's level 9 (third grade):

"Wilbur wants to write a story. He closes the door of his room. Then he sits down at his desk. He gets some paper out of his notebook. Then he picks up his pencil.

5. What will happen next?
a. Wilbur will draw on the door.
b. Wilbur will leave the room.
c. Wilbur will start writing the story."

"To produce correct answers the pupils must accept such test items as a special form of language and then deal with the pragmatics of a situation in which none of the answers are wrong but only one of them is right" (p. 92).

Non-Profit Doesn't Mean Poor

In 1976, *Forbes Magazine* published an article titled "The Pleasures of Non-Profitability" which assessed the Educational Testing Service as an "aggressive and even dynamic growth business." A quick glance at ETS's recent Annual Reports and "990 Tax Returns" demonstrates that the *Forbes* prediction was understated. Today ETS does close to $200,000,000 in business each year with products ranging from the National Assessment of Educational Progress given in public schools across the country to the National Teachers Exam required by nearly half the states. The company's revenues continue to increase by more than 10% annually.

Though ETS now spends more than $100 million a year on salaries alone, all this explosive growth has not gone back into program costs. From 1985-87, revenues greatly exceeded expenses for a total "unprofit" of about $31,000,000. The firm's net worth is now over $68 million.

Non-profitability has also been extraordinarily pleasurable at the individual level. Total compensation to ETS President Gregory Anrig in 1987 was close to $180,000, far in excess of most college presidents. At least a dozen other ETS employees receive more than $100,000 each in salaries and benefits.

Maintaining ETS' privileged position isn't cheap. Corporate legal fees alone come to over $1.5 million. Direct lobbying expenses to prevent public oversight and control adds a quarter million more. All this money, of course, comes directly out of the pockets of test-takers.

TASP *(from page 3)*

well as it does those of Anglo students."

For TASP to have its intended effect of targeting help to those in need, the Texas legislature must appropriate at least $36 million more for remedial programs to serve the 50,000 or more first year students expected to fail the TASP exam. This additional appropriation is not at all certain.

Legislation has been filed to require allocation of the remedial money. The proposal would make the TASP exam one of several criteria, along with, for example, high school and college grades, for entrance to upper level work. It would also prohibit the exam from being more than 50% of a so-called competency index. MALDEF's Kauffman said that the legislation has a good chance of passing this year, due to concern about tests among Black and Hispanic legislators.

FairTest Calendar

FairTest staff members can be heard speaking at the following events:

May	
10	Chapter I Parents, City-wide conference, Detroit, Michigan
13	Appropriate Literacy Evaluation, New York City (*see article, p. 6*)
19-21	"School Reform & the Courts...", sponsored by the ACLU, NAACP LDF and Center for Law & Social Policy, Oak Brook, Illinois
June	
1	New England Association of College Admissions Counselors, Northfield, VT
16	American Assn. of University Professors Annual Meeting, Washington, DC
30	Palm Beach County Schools, Florida

Ann Gould, Community for Effective School Evaluation in Arizona, at the Southern Conference.

Southern Conference

(from page 16)

tion to win a moratorium on testing in grades K-3. This change can be implemented by the State Education Department and will not require legislation.

A Detroit, Michigan group composed primarily of Chapter I parents and teachers also came to Atlanta. Their immediate goal is to modify the use of testing in Chapter I programs. A longer-term goal will be reform of the state's mandatory testing program.

Individuals from Georgia, Mississippi, Louisiana, Tennessee, Oklahoma and Texas also expressed plans to be active on the state or local level in developing testing reform campaigns. The Southern Network on Testing Reform will meet informally by holding fairtesting caucuses at other conferences and by regularly circulating information among its members. To become part of the Network, send your name and address to FairTest. We would like to thank the Mary Reynolds Babcock Foundation for post-conference support.

The New York Times

Copyright © 1989 The New York Times

NEW YORK, WEDNESDAY, JANUARY 18, 1989

Lessons

Edward B. Fiske

Questioning an American rite of passage: How valuable is the S.A.T.?

Most societies have some sort of rite of passage from childhood to adulthood. The aborigine in Australia send adolescents into the wilderness on a "walkabout." German noblemen prove their maturity by acquiring dueling scars.

Americans have one in which their young people spend three high-pressured hours filling in little boxes. Anthropologists call it "Taking the Scholastic Aptitude Test."

Word that the College Board is considering major changes in the test — and that the competiting American College Testing Program has already done so — begs some basic questions about college admissions tests.

First, what do they measure? To anyone with lingering faith in Santa Claus, the Easter Bunny and truth-in-packaging, the answer is obvious: The Scholastic Aptitude Test measures scholastic aptitude. For that answer score zero.

The test was created in 1926, when people still believed that there was something called intelligence that experts could measure and put a number on. Scholastic aptitude was its academic counterpart.

Today no one believes that intelligence is that simple, and developers of the aptitude test have backed away from claiming to measure aptitude. Instead, they say they measure "developed ability," suggesting that something other than genetics is at work in academic success.

The next logical step, of course, would be to change the name of the test — or at least just use the disincarnate letters SAT, like USX. But traditions die hard when they concern rites of passage.

■

What does the test tell colleges? The College Board says the examination predicts college grades in the freshman year. The technical term is predictive validity, which is measured on a scale of zero to one. Zero means that the result would be as random as using a zip code. A perfect 1.0 means that everyone with high test scores will get high grades, everyone with low scores will get low grades and so forth.

Testmakers' research shows that college freshmen grades can be predicted with a validity of 0.42 based on the aptitude test, 0.48 based on high school grades and 0.55 if both are used. Since colleges already have students' grades, the incremental advantage of having the test scores is 0.07. To put this in some perspective, if you know the height of the parent of the same sex, you can predict the height of a child with a validity of 0.51.

This sounds pretty impressive until one goes beyond the numbers.

The test is better at predicting freshmen performance at colleges that are non-selective, small, private, liberal arts, predominantly white and in the Midwest than it is at predicting freshmen performance at colleges that are selective, big, public, technical, predominantly minority and in the Northeast.

When talk turns to specific students the predictive value of admissions tests goes into a nosedive. The statistical error built into the test is such that, of the 1.2 million students who take it each year, 400,000 get scores that are at least 32 points higher or lower than their "true" score, while 120,000 get scores that are off by 64 points or more. This does not include the potential effects of headaches or no sleep the night before.

Do test scores change admissions decisions? In a new book, "The Case Against the S.A.T.," James Crouse and Dale Trusheim of the University of Delaware argue that the test is "valid but redundant." They conducted their own studies and concluded that by using high school grades alone, they could come to the same conclusion as admissions directors in all but 3 percent of cases — not enough, in their opinion, to justify the whole enterprise.

The College Board disputes such assertions but it also argues that the case for its test does not hinge on its predictive value or how many decisions it changes.

"Test scores help students decide where to apply, and they help admissions directors evaluate grade point averages from very different schools," said Robert Seaver, a vice president at the College Board. "Don't colleges have the right to ask for independent confirmation that the kid can do the work?"

Women as a group do slightly better in college than their test scores would predict. Michael Behnke, director of admissions at the Massachusetts Institute of Technology, says that his admissions committee mentally adjusts women's scores upward.

One way to answer the usefulness question would be to require colleges that use the scores to pay part of the cost of the enterprise, but when it comes to the realities of the marketplace, the sponsors are no dummies.

So the debate goes on between the testmakers and their critics. Conspicuously missing are the students who pay $14 for the privilege of showing off their developed ability. At least they have the consolation that filling in little boxes for three hours is better than scrounging for insects in the wilderness or getting your face scarred in a duel.

BEST COPY AVAILABLE

Subscribe to FairTest Publications

FairTest publications will keep you up-to-date on the growing movement for fair & open testing. You will be able to do more than just read about FairTest campaigns —you can help make them successful!

_____ *FairTest Examiner*, published quarterly...Individuals $15/year; Institutional checks $25

_____ *Fallout from the Testing Explosion: How 100 Million Standardized Exams Undermine Equity and Excellence in America's Public Schools*, by Noe Medina and Monty Neill with FairTest staff (45 pages)...$8.95

_____ *Sex Bias in College Admissions Tests: Why Women Lose Out*, by Phyllis Rosser and FairTest staff, 2nd edition (40 pages)...$7.95

_____ *The Reign of ETS*, by Allan Nairn and Ralph Nader (550 pages)...$30

_____ *Beyond Standardized Tests: Admissions Alternatives that Work*, by Amy Allina with FairTest staff (15 pages)...$5.50

_____ *None of the Above: Behind the Myth of Scholastic Aptitude*, by David Owen (327 pp. lists at $7.95)...$5.95.

Updates, issued bi-annually (spring & fall). $7/year/topic:

_____ *Testing Reform Initiatives*
_____ *K-12th Grade Testing*
_____ *University Admissions Tests*
_____ *Teacher Testing*
_____ *Employment Testing*
_____ *Sex Bias in Testing*
_____ *Test Coaching*

Order total _____

Tax-deductible contribution _____

TOTAL ENCLOSED _____
(enclose check to FairTest)

_____ Back issues of *FairTest Examiner* $4 each: Vol___, No.___ (all 8 for $28.00)

(Prepayment on orders under $25.00 is greatly appreciated.)

_____ **Moving.** I have moved/will be moving. My new address is (Please enclose old label or write old address.):

Name ______________________ Title ______________________

Organization ______________________ Phone No. ______________________

Address ______________________ Zip ______________________

Make checks payable to FairTest and mail to: FairTest, 342 Broadway, Cambridge, MA 02139

FairTest Conference Spurs Testing Reform Activity

Nearly 40 educators, children's rights and civil rights activists met in Atlanta, Georgia, on March 18 and 19, 1989, for the first Southern Regional Conference on Testing Reform in the Public Schools. The event, organized by FairTest, was designed to help people plan reform campaigns in their states and localities. It was viewed by all participants as extremely successful. Most committed themselves to working on testing reform issues and agreed to communicate with each other through a Southern Network on Testing Reform.

Other co-sponsors of the Conference were the Alabama Association of Early Childhood Teacher Educators; Alabama Association for Young Children; Atlantic Center for Research in Education (ACRE), North Carolina; Robert Brown, Macon, Georgia; Children's Defense Fund Mississippi Project; Dr. Asa Hilliard, Georgia State University; Intercultural Development Research Association (IDRA), San Antonio, Texas; Dr. Constance Kamii, University of Alabama at Birmingham; Kentucky Association on Children Under Six; Kentucky Youth Advocates; Louisiana Association for the Education of Young Children; National Coalition of Title I Chapter I Parents; and the Oklahoma Association on Children Under Six.

Several speakers described the first wave of success in testing reform in the schools: curtailing the testing of young children. Page McCullough, former Director of ACRE, and Ann Gould of the Community for Effective Student Evaluation (CESE) in Arizona, described how organized efforts altered their state's mandated testing requirements. In North Carolina, the state banned the use of standardized achievement tests in grades 1 and 2, and in Arizona the state stopped requiring the testing of students in grades 1 and 12. (See *FairTest Examiner* Vol. 2, #'s. 2 and 3.)

They were joined by Blair Horner of the New York Public Interest Research Group (NYPIRG) who discussed the various campaigns that NYPIRG has led in New York. NYPIRG is planning to begin work on testing reform in the schools to complement its successful work on university admissions tests.

Another highlight of the conference was the presentation of alternative assessment concepts and practices by Prof. Constance Kamii and Norma Kimzey, the incoming president of the North Carolina Association for the Education of Young Children. Conference participants emphasized the importance of having appropriate assessments available to replace standardized tests.

The examples of successful campaigns and alternative assessments were both inspirational and helpful for planning other campaigns. The participants from Alabama, for example, agreed to organize a statewide meeting by early May to focus on developing strategies for testing reform. Members of the Alabama Association of Early Childhood Teacher Educators and the Alabama Association for Young Children plan to develop a coali-

(continued on page 13)

FairTest
National Center for Fair & Open Testing
342 Broadway
Cambridge, MA 02139
(617) 864-4810

Address Correction Requested

Non-Profit
U.S. Postage
PAID
Cambridge, MA
Permit No. 53278

Please subscribe today.
See page 15.

FairTest

EXAMINER

The National Center for Fair & Open Testing Volume 3, Number 3 Summer 1989

States Move on Testing Reform

While the misuse and overuse of standardized testing is still on the rise a number of state legislatures are taking steps to reverse this trend. In the closing weeks of legislative sessions in Texas, New York, Rhode Island, Massachusetts and Wisconsin, a variety of testing reform bills were considered. They ranged from efforts to curtail testing in the public schools, to bills to protect the rights of university applicants and prevent racial and gender exclusion by tests.

Across the country, testing reform is gaining momentum as more legislatures scrutinize the examiners and their examinations. Below is a description of these efforts:

Texas Drops Grade 1 Test

Texas recently joined the growing list of states that have stopped administering standardized achievement tests to young children. In June, Texas Governor William Clements signed legislation eliminating the first grade Texas Educational Assessment of Minimal Skills (TEAMS) test.

Through this action, Texas joins North Carolina, which banned use of achievement tests in grades 1 and 2 last year and is implementing authentic assessments instead; Mississippi and Georgia, which eliminated state mandated kindergarten tests this year; and Arizona, which stopped requiring its grade 1 test a year ago.

The push to end the grade 1 test had wide support in Texas. The test was seen as providing little useful information while causing undue harm to children.

As part of the same legislation, however, Texas added tests in science and social studies in other grades subject to TEAMS: grades 3, 5, 7, 9, and 11. Proponents of the expansion argued that by testing only reading, arithmetic and writing, TEAMS was narrowing the curriculum in the schools to just those subjects.

New York State Senator Kenneth LaValle (R-Port Jefferson), 1989 sponsor of legislation to establish a "Test-Takers' Bill of Rights," and original sponsor of Truth-in-Testing in New York in 1979.

New York Senate Passes "Bill of Rights"

Following a heated debate that proponents likened to the fight over Truth- in-Testing a decade ago, the New York Senate approved landmark legislation establishing a "Test-Takers' Bill of Rights" for university admissions tests. Final Senate passage came on a 38-17 vote. In the State Assembly, the bill won favorable approval from the Higher Education Committee but did not reach the floor before the legislature's summer recess.

Filed by Sen. Kenneth LaValle, the "father of Truth-in-Testing," the Test-Takers' Bill of Rights creates a series of

(continued on page 13)

CONTENTS

SAT Gender Gap 3
Biased SAT Questions 4
NMSC Results 5
Tests Losing Popularity? 6
Test Criticized in Louisiana 7
Task Force Opposes Tests 8
NY Cancels Regents Exam 8
New Chapter I Regs. 9
Testing in India 10
Wards Cove and Testing 11
Validity Generalization 11
FairTest Book 16

TM017189

FairTest
National Center for
Fair & Open Testing
342 Broadway
Cambridge, MA 02139-1802
(617) 864-4810
(617) 497-2224 (FAX)

Staff: Mike Friedman, Publications Director/Office Manager; Michael King, Summer Intern; Monty Neill, Associate Director; David Riker, Office Assistant; Bob Schaeffer, Communications; Cinthia Schuman, Executive Director; Sarah Stockwell, University Testing Coordinator.

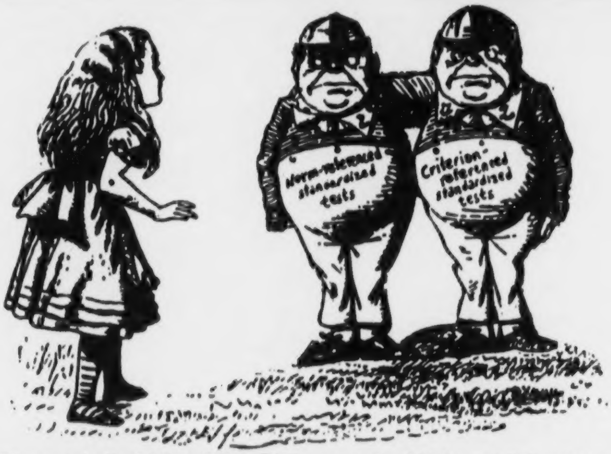
Board of Directors: Denise Carty-Bennia, Esq., Northeastern University Law School, Chair; Dr. Pamela George, North Carolina Central University; Paul Pottinger, Goodsmith and Co.; William Robinson, Esq., Dean, D.C. School of Law; Bob Schaeffer, Treasurer; Steve Solomon, Esq.; Chuck Stone, *Philadelphia Daily News;* Ron Vera, Esq., Barbosa & Vera; John Weiss, Public Policy Initiatives.

National Policy Panel: Leonard Beckum, City University of New York; Dr. Beverly Cole, NAACP; Dr. Harold Dent, Association of Black Psychologists; Arnold Fege, National PTA; Dr. Howard Gardner, Harvard University; Dr. Asa Hilliard, Georgia State University; Rep. Matthew Martinez (D-CA); Deborah Meier, Central Park East Secondary School; Sara Melendez, University of Bridgeport; Aurelio Montemayor, Intercultural Development Research Association; Harry Pachon, National Association of Latino Elected and Appointed Officials; Alfredo de los Santos, Maricopa Community Colleges; Eleanor Smeal, Fund for a Feminist Majority; Theodore Shaw, Esq., NAACP Legal Defense Fund; Dr. Gail Thomas, Texas A & M University; Rep. Ted Weiss (D-NY).

Partial funding for FairTest's work comes from the Rockefeller Family Fund, the Ford Foundation and the Phillip M. Stern Family Fund

Fax it to FairTest

FairTest has joined the facsimile revolution. We can now be reached by fax at (617) 497-2224. Please send us copies of timely news clips and reports.


Norm-referenced standardized tests
Criterion-referenced standardized tests

"Contrariwise," said Tweedledee, "If it was so, it might be; and if it were so, it would be; but as it isn't, it ain't. That's the logic of tests."

Speaking Engagements

FairTest staff members can be heard speaking at the following meetings in the next several months:

August 4-7, "Educational Activists Conference," at University of Cincinnati.

September 27-Oct. 1, National Coalition of Title I/Chapter I Parents, Long Beach, California.

October 5-8, National Association of College Admissions Counselors, New York City.

October 13, National Women's Law Center and National Commission on Testing and Public Policy, Forum on Gender Bias in Testing, Washington, D.C. (see story, p. 3)

November 2-5, National Association for the Education of Young Children, Atlanta.

Quotes of the Month

"The problem can be cast in broader moral terms: the standardized test is disrespectful by design. Mass testing as we know it treats students as objects--as if their education and thought processes were similar and as if the reasons for their answers were irrelevant. Test-takers are not, therefore, treated as *human* subjects whose feedback is essential to the accuracy of the assessment. Pilot standardized tests catch many technical defects in test questions. However, responses to higher-order questions are inherently unpredictable.

The standardized test is thus inherently inequitable . . . inherently unable to encompass the inevitable idiosyncratic cases for which we ought always to make exceptions to the rule."

--Grant Wiggins, *Phi Delta Kappan*, May 1989.

"Current testing programs do not do an adequate job of assessing student thinking. Thinking is a process that is not very adaptable to the multiple choice questions found on most standardized tests. The current testing programs best serve to provide information to those removed from the setting in which learning occurs, to screen, sort and categorize individuals or institutions."

--Kenneth Haskins, former head of the Harvard University Principals' Center, quoted in *ACSD Update*, March 1988.

University Testing

SAT Gender Gap: New Findings Show Causes and Consequences

The vast majority of SAT questions which exhibit large gender differences in correct answer rates are biased in favor of boys, despite girls' superior academic work, according to a new study done by Phyllis Rosser for the Center for Women Policy Studies. *The SAT Gender Gap: Identifying the Causes* includes two item analyses broken down by gender and other factors. One studied an SAT given to 1112 Princeton Review coaching students; the other analyzed ETS data from an SAT taken by 100,000 test-takers in November 1987.

Rosser found that up to 16 percent of SAT questions exhibit a gap of ten percent or more in the correct answer rates of males and females. On the first study, all ten SAT-Math questions with large gender gaps favored boys, despite the fact that girls earned better high school grades in math.

While many of the questions with large gaps did not exhibit "face bias" (an obvious contextual tilt toward one group or another), those with the largest gaps often dealt with the traditional domains and endeavors of males. For example, one that asks about a basketball team's winning record favored males by 27%. Rosser found that, in general, girls did better on questions about relationships, appearance and the humanities, while boys did better on questions about sports, the physical sciences or business.

A surprising new finding was that the higher the grades, the larger the gender gap. Girls with an A+ GPA averaged 23 points lower on the SAT-Verbal section (9 points more than the Verbal M/F gap) and 60 points lower on the SAT-Math section than boys with the same GPA. The smallest gender gap was found at the lowest GPA level.

According to the College Board, this difference exists because girls are earning their high grades in English, humanities and language courses, which are generally easier, while males earn them in the allegedly tougher math and science courses. Among other things, this does not account for the SAT-V gap, where girls should be expected to excel instead of scoring about ten points lower.

An insidious effect of the score gap is the fact that both boys and girls estimate their abilities in English and math to be closer to their SAT scores than their grade point average. Thus, girls judge themselves to be less able than their grades would indicate, as well as less able than boys. Furthermore, "some young women... with A+ GPAs but lower SAT scores had self-selected themselves out of the elite college pool. They were not planning to apply to the most competitive colleges at the same rate as boys with similar grades. In fact, girls in all 4 GPA areas studied planned to go to slightly less prestigious colleges than boys with equivalent GPAs."

In addition to causing girls to lower their sights, the score gap further disadvantages girls in college admissions, receiving score-based tuition awards such as National Merit Scholarships, and being admitted to some gifted and talented programs.

Like males', females' scores increase with their family income level. But a significant gender gap exists at the highest level. Girls from families earning

(Continued on Page 5)

Phyllis Rosser, author of The SAT Gender Gap.

Gender Bias Hearing to be Held

The National Women's Law Center and the National Commission on Testing and Public Policy are jointly sponsoring a hearing on gender bias in testing. The hearing will be held in Washington, D.C. on October 13, 1989 from 9 AM to 5 PM.

Papers will be presented by Phyllis Rosser, author of the new report *The SAT Gender Gap* and FairTest's report *Sex Bias in College Admissions Tests: Why Women Lose Out*, Sylvia Johnson of Howard University, and Esther Diamond, educational consultant. In a follow-up panel, FairTest's Cinthia Schuman, along with representatives from the Center for Women Policy Studies, the American Civil Liberties Union, and the Educational Testing Service will respond to the papers.

The hearing will be held at the Kimball Conference Center, 1400 16th St., NW, Washington, D.C. Contact Laura Epstein at the National Women's Law Center, (202) 328-5160, for more information.

University Testing

Recent SAT Questions

The following questions are from *The SAT Gender Gap* by Phyllis Rosser (see story, p.3). Each passed through ETS's "sensitivity review" to eliminate unfair questions, yet students who aren't familiar with these culturally-specific terms are at a tremendous disadvantage in seeking the right answer. Rosser's analysis shows that many more questions favor males: on the November 1987 SAT, of 23 questions with a male/female correct answer differential of about 10% or more, 21 favored males, suggesting that ETS' review processes are inadequate to detect bias.

1. Although the undefeated visitors _____ triumphed over their underdog opponents, the game was hardly the _____ sportswriters had predicted.

(A) fortunately... upset
(B) unexpectedly... classic
(C) finally... rout**
(D) easily... stalemate
(E) utterly... mismatch

(November 1987)

25% more males than females answered this question correctly.

2.DIVIDENDS:STOCKHOLDERS::

(A) investments:corporations
(B) purchases:customers
(C) royalties:authors**
(D) taxes:workers
(E) mortgages:homeowners

(November 1987)

This question is, on its face, biased against students from lower-income families. It is also biased against females: 15% more males answered this question correctly than females.

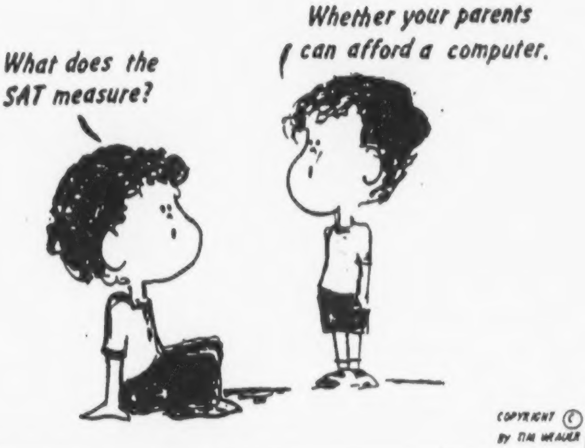
3. PENDANT:JEWELRY::

(A) frame:picture
(B) cue:drama
(C) violin:music
(D) mobile:sculpture**
(E) poetry:prose

(June 1986)

9.6% more females than males answered this question correctly.

** *Wanted Answer*


What does the SAT measure?
Whether your parents can afford a computer.
COPYRIGHT © BY TIM WEAVER

ELECTRONIC EDUCATION VOL. 4 NO. 1/SEPTEMBER 1984

How much does it really cost?

Mrs. Brown's twin daughters, Jane and Julie, are both high school seniors applying to college. How much money will Mrs. Brown send to ETS and the College Board this year?

a) $ 99.62
b) $123.51
c) $185.75
d) $223.91
e) $302.00

Wanted answer: (e)

FairTest calculated that a family with two children taking various tests could easily pay $302.00 to take the SAT ($14.50), get their questions and answers ($10), send scores to colleges ($5.00/school after 3 schools), and take Advanced Placement ($62.00/ea) and Achievement Tests ($21 for up to 3).

New Byrd Scholarships Likely to Be Test-Free

According to the U.S. Department of Education, college admissions test scores are unlikely to be acceptable as the sole criterion for selecting winners for the new Robert C. Byrd Honors Scholarship Program. Final rules for states to follow in administering the federally-funded award were recently made public. FairTest's input to the regulatory process resulted in Secretary of Education Cavazos writing in the *Federal Register* that a state using test scores "would have to be able to show that the test had been designed to test prior academic achievement and not just to predict college performance."

University Testing

Boys Still Receive Lion's Share of 1989 National Merit Awards

The National Merit Scholarship Corporation's (NMSC) exclusive reliance on Preliminary Scholastic Aptitude Test (PSAT) scores to select scholarship semi-finalists continues to be one of the most glaring misuses of university admissions tests. This year, nearly two-thirds of the National Merit Scholarship semi-finalist pool was male, according to a state-by-state analysis conducted by FairTest. This disparity is due to sex bias in the test used to determine eligibility for the awards.

Females totalled only 32% of those eligible for awards, despite the fact that they received better grades in high school. The gender of the remaining students could not be determined from their names. Though NMSC refuses to release the number of winners from minority groups, FairTest estimates an overwhelming percentage are white, as in previous years.

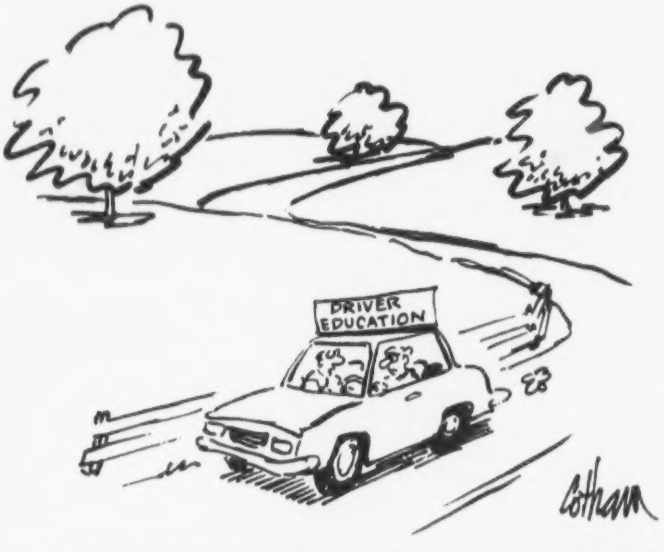
The combined race and gender bias in the test creates a form of "double jeopardy" for minority women, who likely receive a minuscule number of awards. Until National Merit changes its formula, women and minorities will be forced to compete on a playing field that has been proven to be tilted against them.

In New York, a new formula based on grades and SAT scores was developed to award Regents and Empire State scholarships this year, following a federal court decision (see *Examiner*, Winter, 1989).

As a result, girls received $2.8 million more in Empire and Regents scholarships than in 1988 when the SAT was used alone. The New York Public Interest Research Group (NYPIRG) estimates that girls would have won an even greater share of the awards if New York education officials had not weighted SAT results so heavily in their formula.

1989 National Merit Scholarship Semi-Finalists by Gender

Male	63%
Female	32%
Unknown	5%



"Say your car gets 23 miles per gallon at 55 miles per hour and your fuel tank holds 16 gallons. If gas costs 86¢ per gallon and you know the capitals of all 50 states, how much will it cost to drive to the capital of Nebraska?"

Rosser Report

(from page 3)

over $70,000 had lower average scores than boys from families in the $40,000-$50,000 income range. When the test-makers claim the reason for the gender gap is that more lower income girls are taking the test, they ignore the fact that girls score worse than boys at every income level.

Rosser found the timed nature of the test to be a factor in girls' lower SAT-M scores. (This is true for Blacks and Hispanics as well). She points out that "[t]his highly speeded test rewards the facile test taker rather than the sophisticated, thoughtful thinker who gathers new information and organizes, evaluates, and expresses original thoughts clearly and concisely."

Among Rosser's recommendations are: ETS should eliminate SAT questions which show the largest gender, race, and class differences; the SAT-V should be equalized so that men and women score equally well; test makers should publicize their validity research to the public; test makers should do further research on individual questions; and the SAT should be less speeded.

The SAT Gender Gap: Identifying the Causes, is available from the Center for Women Policy Studies, 2000 P St., NW, Suite 508, Washington, DC 20036, for $15.00.

Tests Losing Popularity?

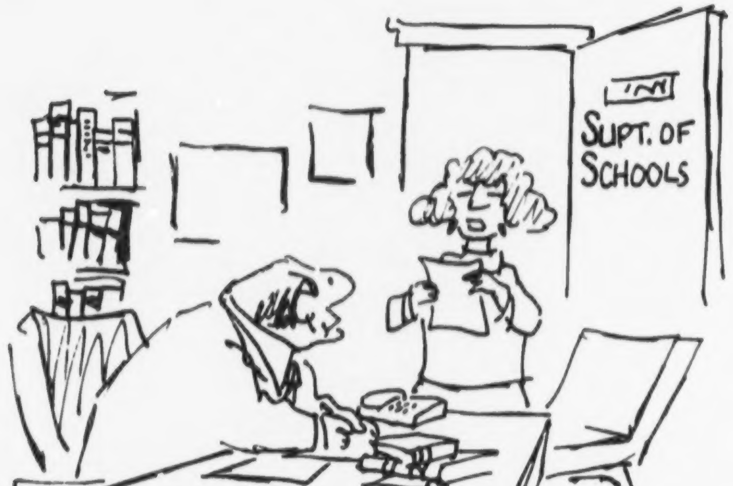
Two recent national polls suggest that the public's apparent love affair with standardized testing may be on the wane. In the surveys, clear majorities indicated disbelief about the accuracy of testing in education and for employment.

A Media General/Associated Press poll on public education, conducted in May 1989, asked, "Do you think standardized test scores reflect how well students are learning in school, or not?" Of 1084 respondents, 54% said "No," 37% said "Yes," and 9% either did not know or did not respond.

In almost every population sub-group, the "No" votes outnumbered the "Yes" votes. Hispanics were most likely to say "No," though clear majorities of whites and blacks also said that tests do not reflect how well students are learning. Women were more likely to say "No" than were men; and younger respondents selected "No" more often than did older respondents.

A January 1989 Harris poll conducted for the NAACP Legal Defense and Educational Fund produced similar results about employment tests. It asked, "Do you feel that giving standardized written tests to people before they are hired accurately determines their competence to perform a particular job, or not?" Of 2103 respondents, only 31% agreed that tests "Accurately determine competence," while 66% believed they "Do not accurately determine competence" and 3% were "Not sure."

A majority in every sub-category believed tests were not accurate. In this poll, whites were slightly more distrustful of tests than blacks, Hispanics or Asians, and men believed the tests were inaccurate more than did women. Again younger respondents disbelieved test results more than did older respondents.



"There's been a slight mistake. 1542 wasn't the average SAT score— it was the number of students who took the test."

Students Sabotage Test in California

In a rebellion against standardized exam overkill, students at West High School in Torrance, California, sabotaged their answers on a recent California Assessment Program (CAP) test. Average scores on the reading test dropped from the 85th percentile to the 51st percentile, and in math fell from the 95th to the 71st percentile. Some test booklets included no correct answers.

Some students were angered by teachers interrupting classes to prepare students for the test and by administrators emphasizing the importance of the test.

Kelle Price, president of the student body, reported that some students were angered by teachers interrupting classes to prepare students for the test and by administrators emphasizing the importance of the test.

Two years ago, California students in at least one school threatened to boycott the CAP unless they could decide what to do with the money awarded to schools for good performance on the state-mandated test.

K-12 In Brief

The **Minnesota** legislature has rejected Democratic Governor Rudy Perpich's proposal to implement a state-wide testing program. However, the legislature agreed to fund 10 sites where "learner outcomes" will be researched. The sites may choose multiple forms of outcomes assessments.

The recently passed **Oklahoma** Education Challenge 2000 Act expands the state's student testing program. Among the new tests will be a high-school graduation competency exam, to take effect in 1993.

Grade School Testing

Black Task Force Criticizes Testing

"[O]ver-emphasis upon standardized testing subverts true education, undermining the curriculum and eroding the quality of teaching," argues a committee of the Joint Center for Political Studies in *Visions of a Better Way: A Black Appraisal of Public Schooling.* As a result of this and other shortcomings, the report says, "[S]erious questions must be raised about the validity of standardized testing and its effects not only upon black and minority children but upon quality education for all. We advocate the development and sensitive use of a variety of methods for assessing both school and student performance" (p. 19).

The criticism of testing is one part of a comprehensive study of public schools that includes recommendations for reform. The Committee was chaired by Dr. John Hope Franklin and included some 30 prominent African-Americans. The report is available from the Joint Center for Political Studies, 1301 Pennsylvania Ave., NW, Suite 400, Washington, D.C. 20004 for $9.00.

Controversies Surround Louisiana Test

Louisiana's new high school graduation exam has become embroiled in controversy in its very first year. Private schools already have won an exemption from the test, the state lacks the funds for remediation for those who fail, and concern has arisen because essays are being secretly flagged for signs of problems. The controversies have provoked several unsuccessful efforts to block the exam.

The test, which high school students must pass to graduate starting with the class of 1991, includes 10th-grade reading, writing and math as well as 11th-grade science and social studies exams.

Responding to political pressure from parochial schools, the legislature passed and Governor Buddy Roemer signed an exemption from the test for private schools. Roemer, who signed it even though he supports testing all students, suggested that private schools may need a somewhat different test because their curriculum does not match that of the public schools. Although one of the test's proponents vowed, "Everybody is going to take the test or nobody is going to take the test," his motion to stop all testing failed.

Meanwhile, concerns have been raised as to what will happen to the more than 25% of students who have failed one or more parts of the exam. Remedial classes are supposed to be offered, but only $5.2 million is available out of the estimated $12.5 million needed.

"We're all a bunch of hypocrites to endorse a testing program that identifies endangered students but fails to provide money to help them succeed," said State Rep. Joe Delpit. His effort to stop the testing on those grounds also failed. Supporters of efforts to drop the test also argued that it would penalize students from low-income school districts and force up the drop-out rate.

In a third controversy, the essay writing portion of the exam is being flagged when the reader suspects the author may be suicidal, have a drug dependency, suffer from child abuse, or have some other problem. The graders, employed by Measurement, Inc., of North Carolina, are teachers without training in counseling.

The flagging caused widespread protest as an invasion of privacy. Neither parents nor students were informed the papers would be flagged. Defenders of the process argue they are under a legal obligation to identify children who may be in trouble.

Measurement, Inc. reported that it flags questionable papers by students from other states, including Mississippi, New Jersey, South Carolina and Texas. The Louisiana State Board of Elementary and Secondary Education has decided to continue the practice.

(Source: *Baton Rouge Morning Advocate*)

Suggested Reading

Educational Leadership for April 1989 is devoted to "Redirecting Assessment." It contains 17 articles covering issues from why assessment must change to differing proposals for appropriate and authentic alternatives. It is a thorough introduction to the problems and solutions concerning standardized testing. (Order Processing, ASCD, 1250 N. Pitt St., Alexandria, VA 22314; $5 pre-paid.)

Right from the Start, a report from the National Association of State Boards of Education on early childhood education, contains some strong recommendations on testing. For example, it urges that no major decision be based on a single test score. Test results should not be used to place or retain children or determine eligibility for enrollment. In general, assessment ought to flow from and reinforce a developmentally appropriate curriculum. The curriculum should not be a worksheet, drill-and-practice regimen driven by standardized testing. This report would make a fine gift for a school board member, administrator, teacher or active parent. (Available from NASBE, 1012 Cameron St., Alexandria, VA 22314; $5.)

Grade School Testing

Milwaukee Task Force Opposes Tests

The Milwaukee Public Schools Assessment Task Force has called for a five-year plan "for the development of authentic assessments to replace the norm-referenced tests currently used." The draft report, dated June 2, 1989, criticizes most testing for failing "to measure meaningful knowledge and competence" and contributing to "inequity in educational opportunity." Moreover, standardized testing is rarely used to improve education for students, is frequently inappropriate, provides only a limited picture of a child's abilities, and fails to offer useful information to parents.

94,000 Students, 480,000 Tests

The Task Force found that, in Milwaukee, some 94,000 students are administered over 480,000 standardized tests each year, or about five per year per student. This does not include the estimated 300,000 standardized basal tests also administered annually. This figure is 5-8 times larger than the number Milwaukee school officials reported to FairTest for *Fallout from the Testing Explosion*.

Curricular Outcomes/Authentic Tasks

The Task Force recommends "that all assessment be linked to the curricular outcomes determined by the district through the design of authentic tasks," that the district commit the resources required to develop the new assessments, and that the school board "support legislative initiatives to reduce state mandated testing" in special program schools.

"We believe," concludes the report, "that unless assessment is grounded in curriculum development and directed to the improvement of each student, we cannot achieve meaningful accountability."

The report discusses each criticism and recommendation in detail, providing a rich analysis of the problems of testing in the Milwaukee schools as well as suggestions for initiating authentic assessments. Reports from two other task forces, on curriculum and urban youth, also called for changes in curriculum and assessment to avoid standardized test-driven instruction. The school board has accepted the three reports and will begin to act on them this summer.

> *"...unless assessment is grounded in curriculum development and directed to the improvement of each student, we cannot achieve meaningful accountability."*
>
> *--Milwaukee Public Schools Assessment Task Force*

Major Change in Assessment

If implemented, the Task Force's recommendations would produce a major change in the nature and role of assessment in Milwaukee's schools. The report is a model for other local activists and school officials interested in studying current practices and moving toward implementing appropriate assessments.

(A discussion of testing in Milwaukee and the Task Force recommendations can be found in the May-June 1989 *Rethinking Schools*, available from *Rethinking Schools*, Box 93371, Milwaukee, WI 53203; $2.)

New York Cancels Regents

Following the printing of the answer key in the *New York Post* and the discovery that several different exams and answer keys had been stolen, the New York State Department of Education (SED) cancelled its June administration of the Regent's exam in Chemistry, which is taken by thousands of high school students each year. The SED has directed each school district to investigate all other Regent's exams given in June for evidence that students had access to test forms or answers. If widespread cheating is discovered, the scores will be cancelled.

The *Post* printed the answer key to one form of the Chemistry exam on the front page of its June 20 Metro edition under the headline "Easy as Pi." The *Post* said that its reporters had obtained several copies of the exam after only 15 minutes and two phone calls. The paper also reported that some students were paying anywhere from $10 to $2000 for copies of the exam.

According to students quoted by the *Post*, "Getting the tests is a breeze. If the school boards or anyone else thinks this is an isolated thing, they're wrong." Other students boasted that they had been able to buy keys to locked vaults and even the principal's office from janitors. They then were able to dismantle the metal box the tests were shipped in, slip one test and answer key out, copy it and return it. "It can be made to look like it was damaged on the way from Albany," one student explained.

Scandals such as the one in Los Angeles last year (see *FairTest Examiner*, Vol. 2, #4) and now the "Examscam" in New York are just the latest in a long line of cheating incidents that stem from the undue emphasis that is placed on test scores. Instead of reducing testing, though, the state's response to the cheating has been to emphasize the need for increased test security.

Grade School Testing

New Chapter I Regulations

The National Coalition of Title I/ Chapter I Parents and the National Congress of Parents and Teachers have criticized new Chapter I regulations for relying too heavily on standardized testing. Chapter I is the nation's primary federally-funded remedial education program.

Annual Norm-Referenced Testing

The new federal regulations for Chapter I require that each student be tested with a norm referenced instrument once a year in each Ch. I subject in which s/he is enrolled (i.e., reading and/or math). However, the regulations leave open the possibility that, aside from the yearly requirement, more appropriate assessments can be introduced.

Parents and teachers could use the increased power they have under the new Act to insist that Ch. I programs be revised to include comprehensive, locally determined goals that are assessed through appropriate instruments, not multiple-choice tests.

Using Alternatives to Tests

Under the regulations, local districts are to determine precise curriculum content and the measures of student progress. The states, however, may impose requirements beyond those in the federal regulations. These could be test-score increases, or they could be measures other than test scores.

Local programs could easily meet the federally mandated gains without making tests the instructional focus, particularly if the states have not imposed higher testing requirements. Localities could, in other words, ignore the tests and concentrate on alternative measures.

Three Years to Catch Up

According to the Act, students in Ch. I programs are to receive instruction that includes both basic and higher order skills. Progress is to be measured in both skill areas, and alternatives to standardized exams are possible.

Also according to the law, students in the program should make sufficient progress to catch up with their non-Ch. I peers within three years. This requirement is absent from the federal regulations. Rather, the regulations only require that students "make gains beyond that which they would be expected to make in the absence of the additional help the program provided." Minimal test score gains would meet this criterion.

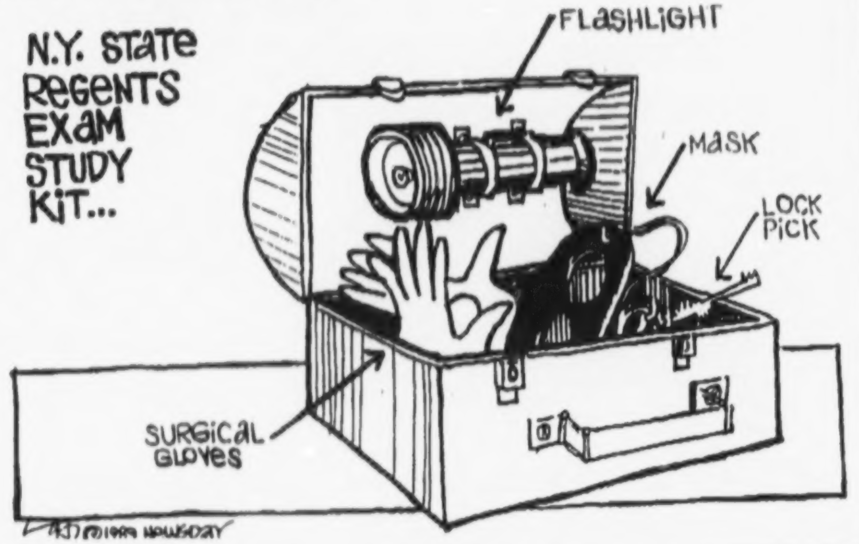
The Ch. I Parents and the PTA criticized this "improvement" requirement. They argue that progress should be measured as closing the gap with other students. Too often, they claim, Ch. I students fall further behind their peers.

What local programs actually do largely will depend on the actions of parents, teachers and others who want a rich and full Ch. I program. Authentic assessment could be developed to define and measure locally determined program outcomes in both basic and higher order skills and in closing the gap between Ch. I and other students.

(FairTest's Monty Neill, with Susan Harman and Brenda Walcott of the American Reading Council and Carole Edelsky of Arizona State University, will be at the National Conference of Ch. I Parents in September. They will make presentations on testing and alternatives, including workshops on implementing appropriate assessments in Ch. I.

Exam Following Theft

M.G. LORD'S VIEW



Testing in India

Testing in India is used to support the interests of a small minority of the population, to the detriment of the majority of Indians, according to information contained in a packet of material sent to FairTest by M.M. Pichare of Hyderabad, a city in southern India.

Pichare indicated that major tests for university entrance and the civil service are designed by persons drawn from "the urban middle and upper class," and "administered in big cities only." People from these backgrounds "want to perpetuate the present situation by insisting on 'merit' (as measured by the tests) in admissions to educational institutions and in employment. They clamour for the 'objective type' tests." Pichare also reported that coaching for these tests "is flourishing in urban areas," and supplied copies of several coaching school advertisements, all of which claim to offer the best chance of getting into the most prestigious universities.

Mr. Pichare's documents indicated that upper caste people, who make up only 24% of the Hindu population of the country, hold 89% of the higher level civil service jobs reserved for Hindus. These jobs all use tests for entry. In addition, Pichare says, "no good predictive validity study of any of these tests has been done. Information about predictive validity of only one test, the National Talent Search Examination, is available publicly...It is almost impossible to get any information about most of the tests used in India."

One sample booklet says in its preface "Objective Type tests are designed to measure not only factual knowledge, information or understanding, but also the higher intellectual abilities such as the ability to apply knowledge to new situations, to analyse and evaluate facts, data, events or situations." All of these factors supposedly can be determined from a 200 question multiple-choice examination. The instructions also caution against guessing, although there is no penalty for wrong answers.

Late last year, Pichare and two colleagues in Hyderabad filed a petition in court questioning the validity of an examination given to award research fellowships. They say that the exam has a disparate effect on students from non-metropolitan universities and rural areas. The matter is still pending in court.

Reprints Available from FairTest

Article Backs *Golden Rule* Settlement

Almost since the day company leaders signed-off on the landmark Golden Rule settlement establishing a new way to make the Illinois Insurance Agent Licensing Exam fairer but not easier, the Educational Testing Service has conducted an aggressive campaign to minimize the agreement's impact.

In the Winter 1989 issue of the *Valparaiso University Law Review*, Golden Rule's attorneys Martin Shapiro, Michael Slutsky and Richard Watt respond most effectively. "Minimizing Unnecessary Differences in Occupational Testing" is a comprehensive review of the court case, test construction techniques, and the reasons why conventional psychometric approaches fail to eliminate biased questions. More than two hundred footnotes provide a wealth of references for anyone involved in research or litigation concerning bias in standardized, multiple-choice exams.

Reprints of this important article are available from FairTest. Please send a large self-addressed envelope (at least 8" x 10") with $1.05 postage to order your copy.

FairTest Testifies at U.S. Civil Rights Commission

FairTest provided testimony to a June 16 consultation of the U.S. Commission on Civil Rights on testing in education and employment. Scheduled as the first in a series of hearings, this session focused on test construction (validity and bias) and legal/policy issues. In testimony submitted in writing, FairTest criticized the inadequate validity of most tests used in schools, college admissions and employment, and offered recommendations for making all assessments fair, open, valid and relevant.

Other presenters included Dr. James Loewen (U. Vermont), Dr. Nancy Cole (Educational Testing Service), Dr. Lloyd Bond (University of N. Carolina-Greensboro), Dr. Alexandra Wigdor (National Research Council), Clint Bolick (Landmark Center for Civil Rights), and Barry Goldstein (NAACP Legal Defense Fund). A transcript of the consultation, including discussions that followed the presentations, will be published by the Commission, along with longer papers to be prepared by each of the presenters.

Literacy Evaluation Conference

On May 13, nearly 100 people met in New York City for a conference on Appropriate Literacy Evaluation. The day-long event, co-sponsored by FairTest, focused on how to reduce reliance on standardized tests, how to implement alternatives, and what good literacy evaluation would be. Participants discussed appropriate assessments for young children, public schools, university admissions and adult literacy programs.

As a result of the conference, groups are working on changing adult literacy evaluation in New York City and on K-12 testing reform in New York State. Additionally, a handbook on appropriate literacy evaluation is being prepared for fall distribution.

Employment Testing

Supreme Court Undermines Employment Test Challenges

In a recent decision, the United States Supreme Court severely limited the ability of minorities and women to legally challenge discriminatory employment practices, including biased employment tests. With the threat of lawsuits substantially reduced, employers are more likely to establish and maintain discriminatory tests.

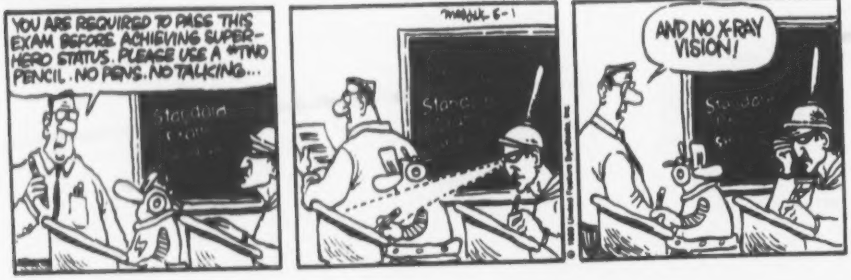
The 5-to-4 opinion in *Wards Cove Packing Company v. Antonio* increased the evidence plaintiffs need to show in order to succeed in certain employment discrimination cases under Title VII of the Civil Rights Act. Reversing precedents dating from a 1971 case, *Griggs v. Duke Power*, the ruling shifted the burden of proving racial and other types of discrimination to the employee. Since employees have limited access to the evidence needed to prove discrimination, this shift has devastating implications for the future of civil rights and women's rights litigation. Previously, employees could prove discrimination using statistical evidence of racial or gender imbalances in the workforce. Then the burden shifted to the employer to prove the "business necessity" of the practices that caused the imbalance.

In *Wards Cove*, however, the Court held that a "mere" showing of statistics will no longer suffice. Instead, plaintiffs must demonstrate that statistical disparities are the result of specific employment practices, each of which must be shown to have a significantly discriminatory effect.

Even if plaintiffs are able to meet this heavier burden, the ruling still makes it much easier for employers to defend their actions. They need only provide evidence that the challenged practice significantly "serves the legitimate employment goals of the employer." This is a further narrowing of *Griggs*, which required employers to prove that employment practices which disproportionately harmed minorities and women were "essential" to the operation of a business.

Because *Griggs* involved a successful challenge to an employment test, its implicit rejection is certain to have a major impact on the future of employment testing litigation, as well as other civil rights claims. Under the *Griggs* standard, lawyers have been able to overturn hundreds of employment tests used by civil service and private employers. *Griggs* also provided guidance for Title VI and Title IX cases, prohibiting race and sex discrimination in federally-funded educational programs. Now, due to the increased risks and costs of such litigation, many lawyers will be discouraged from bringing suits under these statutes.

ROBOTMAN by Jim Meddick


YOU ARE REQUIRED TO PASS THIS EXAM BEFORE ACHIEVING SUPER-HERO STATUS. PLEASE USE A #TWO PENCIL. NO PENS. NO TALKING...
MEDDICK 6-1
AND NO X-RAY VISION!

Federal Tests Questioned

Shortly before the Supreme Court's recent *Wards Cove* ruling, which undermined key job discrimination precedents, a federal commission provided new evidence that employment tests are often unfair and irrelevant. The National Research Council's report *Fairness in Employment Testing: Validity Generalization, Minority Issues and the General Aptitude Test Battery* generally sides with those who criticize the U.S. Employment Service's faith in test scores as a measure of worker ability.

In its most controversial recommendation, the NRC supports "within group scoring" and similar adjustments for black and Hispanic applicants "because the effects of imperfect prediction fall more heavily on minority applicants." The Reagan Administration Justice Department had challenged this practice which requires referral based on an applicant's percentile rank in his or her ethnic group, not on raw scores.

The report also criticizes research endorsing the "validity generalization" concept that one test can predict equally well for all jobs. That concept exaggerates both the predictive accuracy of the General Aptitude Test Battery (GATB) and the economic gains derived from basing job referrals on test scores. It concludes that, "Because tests provide only partial information about future job performance, we recommend that Job Service offices that adopt the VG-GATB Referral System continue to use multiple criteria in choosing which applicants to refer." In addition, it recommends that "no job seeker be obliged to take the GATB; every local office that uses VG-GATB referral should maintain an alternative referral path for those who choose not to take the test."

Copies of the report are available for $29.95 (paperback) from National Academy Press, 2101 Constitution Ave. N.W., Washington D.C. 20418.

The New York Times

NEW YORK, THURSDAY, MAY 11, 1989

Copyright © 1989 The New York Times

Debate Intensifying On Screening Tests Before Kindergarten

By DEIRDRE CARMODY

The screening of children to determine whether they are ready for kindergarten has become the focus of an intense debate among early childhood experts, teachers and parents around the country.

Many educators say misuse of the tests turns some children into failures before they even enter kindergarten. But proponents say the assessments determine a child's developmental age, which they view as a better measure of readiness for kindergarten than chronological age.

While assessments for 4- and 5-year-olds are not new, the concern is that these tests are increasingly being misused as admissions tests and children who do not do well are being forced to put off starting school.

Federal law mandates the developmental screening of young children to identify those with special needs. Most assessment tests are intended to be supplemented by parental information, medical history and evaluation by trained observers. But they are being used more and more as pass-fail examinations.

Typically, parents take their children to the school for a test lasting 20 to 30 minutes. Children are asked questions like "What meows?" and "What

Continued on Page B13, Column 1

Screening for Kindergarten Inspiring U.S. Debate

Continued From Page A1

kinds of things float?" They are asked to copy forms like a circle, a cross and a square. They are given a stick figure of a man to complete, are told to name some animals and are asked to talk about their interests.

Last May Beth Bradley brought her son Shane, 4 years old, to be tested for the Denver City, Tex., kindergarten, as required by the school. In September, Mrs. Bradley and other parents were told that their children had been grouped according to their developmental age into "young 5" instead of kindergarten so that they "would not see other children doing better and feel bad about themselves," she said.

Mrs. Bradley said the school told parents that children in both groups were taught the same things. But when she and other parents monitored the other classes, she said, they found those children who had tested higher were on phonics programs, using computers and doing beginning mathematics.

"Our children, on the other hand, were just beginning to cut pictures out of magazines and do alphabet sounds," she said.

At the end of the year, the school will decide whether Shane will go into regular kindergarten or first grade. "Our school says that Gesell is giving these children the gift of a year; we say it is the theft of time," Mrs. Bradley said. "We have begged our school board that the test be wiped out because this is detrimental to our children."

Principal Defends Screening

Brenda Massey, the principal of the Denver City kindergarten, disagrees.

"It is just an instrument to back up our other observations and we feel that it is telling us what we need to know," she said. "I have taught in situations where there are all different developmental ages in a classroom. If one child was ready to do paper-pencil work and another was not, that child was unhappy. All I have seen here is successful children."

Are children being stigmatized as failures even before school?

No one knows how widespread the misuse of such screening may be, since the practice varies from district to district and in many cases from school to school. Dr. Samuel J. Meisels, professor of education at the Center for Human Growth and Development in Ann Arbor, Mich., says the misuse is "shockingly high."

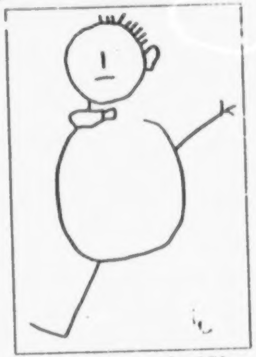
While there are also no overall data on how many children have been held back because of their performance on the tests, Dr. Meisels estimates that more than 20 percent of those enrolled in first grade are at least a year overage for that grade.

Many educators are troubled by the social and educational implications of this. They point out that young children develop at widely differing rates, so that those deemed immature when they are tested in the spring may well have reached the norm by the time school begins in the fall. By that time, of course, it is too late and, in effect, they have lost a year of their lives.

Fear of Long-Term Effect

"Research shows that children who are behind even one year have a significant increase in being at risk of not completing high school; this risk increases to 90 percent when students are two years behind," Lydia Richards, supervisor of quality assurance, division of monitoring and evaluation, in the New York State Department of Education told prekindergarten directors in Albany last month.

Another concern is that putting children with low scores in one group becomes a self-fulfilling prophecy that creates low expectations, a negative self-esteem on the part of the child and the likelihood of repeated low performance.



Part of the Gesell test for children about to enter kindergarten consists of this incompletely drawn man, first used in 1925. The child is asked to complete the figure. The result provides information about eye-hand coordination, awareness of detail and sensitivity to emotional expression.

"Kids who are held back do no better than kids who were not held back," said D. Monty Neill, managing director of FairTest, a Cambridge-based group that monitors tests. "It evens out at the end of third grade, but one group is a year behind their age group. And there have been studies concluding that in general, retention does psychological harm."

Still, Rosalind Charlesworth, professor of curriculum and instruction in the College of Education at Louisiana State University, says that the misuse of kindergarten assessment tests is a subject of wide concern.

"Every time I go to a professional meeting, everybody is going berserk about this subject," she said.

She traces the use of these tests to the philosophy of teaching children earlier and earlier and testing them more.

"You end up with tests forming the curriculum," she said. "Pretty soon the first-grade teachers are saying, 'Oh, they are not coming in prepared.' So then you get into the kindergarten entrance exam routine."

Teaching Clock Is Advanced

Sue Bredekamp, director of professional development at the National Association for the Education of Young Children, the largest professional organization of early childhood educators, says what used to be taught in first grade is now taught in kindergarten. This is why children are being tested beforehand to see if they will be able to handle the material.

"The debate has been over the issue of how we best protect children from inappropriate curriculum practices and expectations in kindergarten and first grade," she said.

In some ways, she said, the tests are actually making the problem worse by screening out the "immature children" and, in effect, giving license to the regular classroom to become more academic and formal.

As with the Scholastic Aptitude Tests, the use of kindergarten placement tests has led to coaching by parents and tutors. It has also led to coach-them-yourself books, like "The Baby Boards, A Parents' Guide to Preschool and Primary School Entrance Tests," published by Arco.

The most commonly used of the assessments was developed by the Gesell Institute of Human Development in New Haven. The assessment, which is used by 15 to 18 percent of the schools in the country, is meant to measure the overall behavioral readiness of children to adjust and adapt to appropriate age and grade expectations.

Need to Complement Test

Many educators reject the test as an inaccurate barometer of development; even proponents say the test should be used in connection with the observations of teachers and trained developmental examiners.

"We do not condone the use of the Gesell assessment as the sole determinant of a child's placement in school," Susan Sweitzer, director of education and training at the Gesell Institute, said. "Parents should always have the final say regarding a child's placement in school."

When a visitor, who had come to observe the testing of prekindergarten children at the Ardsley, N.Y., school system, asked the name of the assessment test, the kindergarten teacher declined to identify it.

"Our parents are very competitive," said the teacher, Stephanie Hamilton. "If you print the name of the test, our parents will get hold of it and prepare their children for the questions."

Ardsley's purpose in giving tests is simply to find out more about the strengths and weaknesses of the children. "We never tell parents to keep their child out, but if there is a special problem we tell the parents about it," Ms. Hamilton said.

Dr. Robert Lichtenstein, whose field is psycho-educational assessment and who is doing an evaluation study of extra year programs, said that he had not quite made up his mind.

"My position is that I am having a hard time with both of the extreme positions that educators are seemingly being forced to choose between," he said.

BEST COPY AVAILABLE

States Move on Testing Reform

(from page 1)

"due process" protections. These new rights include a set of standards to assure that students accused of cheating (like the Mexican-American high schoolers in the movie "Stand and Deliver") are treated as innocent until proven guilty. Other provisions require that test registration books contain additional information about the content and accuracy of the exams as well as details about administration procedures. The proposal would also permit score cancellations if a student systematically mismarks his or her answer sheet by omitting a question without skipping the corresponding answer line.

Most controversial was a provision requiring at least two exams annually to include no experimental section.

Advocates argued that these "pre-test" items are often biased, ambiguous or confusing. Students who react negatively to such flawed questions may panic or become confused. As a result, their test scores can plummet.

Testing companies, led by the Educational Testing Service and the College Board, launched an all-out drive to block the bill, flying high-ranking executives to Albany and hiring the state's top lobbyists.

The more than two-to-one margin of Senate support for the Test-Takers' Bill of Rights, despite the high-pressure campaign against it, demonstrates the power of this issue. Sponsors, led by the New York Public Interest Research Group (NYPIRG), hope that the full Assembly will take up the Test-Takers' Bill of Rights if it returns this fall. Otherwise, the proposal remains alive for consideration early in 1990.

R.I. Teacher Test Moratorium

The Rhode Island Assembly overwhelmingly approved legislation endorsing a moratorium on the use of the National Teacher Examination (NTE) until the state demonstrates that the test accurately measures teacher competence. The legislation was introduced in response to a report by a legislative commission calling for a ban on the NTE (see Spring, 1989 *Examiner*). The bill, introduced by commission member Representative Harold Metts, was strongly backed by the Rhode Island affiliate of the American Civil Liberties Union and other civil rights organizations including the NAACP and the Urban League.

H. 6121 states that "No person shall be barred from certification . . . solely because of his or her score on any standardized teacher examination." The state board of regents may require teacher candidates who fail the test to pass a job performance review after three years in the classroom.

Strong opposition from Rhode Island education officials blocked passage of the plan in the State Senate, but sponsors expect to revive it when the legislature returns next year.

Mass. Scholarship Legislation

Legislation to eliminate the Massachusetts Senatorial Honors Scholarship, which bases its awards solely on SAT scores, has passed the Commonwealth's Joint Committee on Education and is currently pending in the House Ways and Means Committee. In the meantime, the controversy surrounding the scholarship combined with the state's fiscal crisis, has led Governor Dukakis to veto the scholarship appropriation for the current academic year. Sponsors hope the bill to permanently repeal the discriminatory scholarship program will be considered when the legislature returns from its summer recess.

Until recently, Massachusetts and New York were the only states to award college scholarships solely on the basis of SAT scores. In February 1989, a federal district court struck down the New York scholarship program on the grounds that it discriminated against females, who received only 40% of scholarship funds, despite their higher academic achievement. A disproportionate share of the Massachusetts awards also has gone to men, prompting the scholarship reform bill.

Wisconsin to Stop Testing K-2?

Legislation will be filed in late July in Madison to ban standardized testing in Kindergarten through grade 2 and to develop and implement appropriate assessments.

The bill, drafted by Representative Louis Fortas and co-sponsors, requires the state to "develop and make available to school districts developmentally appropriate, individualized assessment instruments for pupils in kindergarten to grade 2." It also says, "A school board may not administer standardized tests to pupils in kindergarten to grade 2," but "may administer" the developmentally appropriate instruments. The *Rethinking Schools* group in Milwaukee is providing technical assistance and support to the legislative campaign.

The Wisconsin bill is modeled on the North Carolina testing ban for grades 1 and 2, but its coverage is broader. Currently, Wisconsin does not mandate mass standardized testing, but many school districts administer such exams.

FairTest Book

(from page 16)

readers at the cover price of $6.95, postpaid. Simply use the form on p. 15 to order your copy today!

Standing Up to the SAT also makes an excellent gift for your school guidance counselor, college admissions officer or local education writer as well as for any high school junior or senior you know. Call FairTest for information on bulk purchases.

Members of the media may obtain review copies by calling ARCO at (212) 373-8946.

The New York Times

Copyright © 1989 The New York Times

NEW YORK, TUESDAY, MAY 23, 1989

Letters

Testers, Not Women, Must Explain S.A.T. Gap

To the Editor:

Gertrude Himmelfarb's defense of the Scholastic Aptitude Test against charges of gender bias ("Self-Defeating Feminism," Op-Ed, May 8), unfortunately, misses the point. The growing concern about the approximately 60-point gender gap in S.A.T. scores stems not from generalized notions of "liberty," but from the concrete fact that the S.A.T. fails to accomplish its stated purpose of accurately predicting first-year college grades for young women. Although females' test scores are significantly below those of males, young women consistently earn the higher grades.

The score differential is not controverted. The principal excuse holds that female test takers are more often from minority groups, from poorer families or have taken fewer advanced courses in high school, and that these — not gender — account for the differential. Yet no matter how the identity of the test takers is broken down and these factors are controlled for — by ethnic group, family income or course preparation, singly or in combination — young men consistently score higher than young women.

The Supreme Court has ruled in Price Waterhouse v. Hopkins (front page, May 2) that once a woman proves gender played a substantial role in an adverse employment action against her, the burden shifts to the employer to justify its actions. The same analysis should apply here. The evidence shows that the S.A.T. fails to achieve its purpose as it underpredicts the first-year college grades of more than 50 percent of the test takers. The victims of the bias should not have the responsibility of explaining the cause of the scoring skew. The Educational Testing Service and the College Board, which prepare and sponsor the S.A.T., must bear the burden of explaining why they do not design a test that accurately predicts college performance for both young men and young women.

The test score gender gap has damaging ramifications. Many colleges and universities rely heavily on S.A.T. scores in their admissions to the detriment of lower-scoring young women. Only one-third of the prestigious and valuable National Merit Scholarships are awarded to women. The Johns Hopkins Center for Academically Talented Youth, a program for gifted teen-agers, bases its admissions solely on the S.A.T. scores of a selected group of seventh graders. While the applicant pool is equally divided between boys and girls, boys fill two-thirds of the slots, based on S.A.T. scores. These are only a few examples.

Remedying the gender bias in the S.A.T. can only increase young women's opportunities comparable with their abilities — a goal consistent with the feminism of yesterday and today.

ELLEN VARGYAS
LAURA EPSTEIN
Washington, May 10, 1989

The writers are, respectively, a staff lawyer and public education specialist, National Women's Law Center.

•

Near Zero Probability

To the Editor:

Gertrude Himmelfarb ("Self-Defeating Feminism," Op-Ed, May 8) misunderstands the criticism of gender bias in the Scholastic Aptitude Test. Even the test's manufacturers agree that the S.A.T. underpredicts female college performance and overpredicts that of males.

This is the conclusion reached by Federal District Judge John M. Walker in a landmark ruling striking down New York State's reliance on S.A.T. scores to award up to $40 million in college scholarships annually (front page, Feb. 4). As Judge Walker said, "After a careful review of the evidence, this court concludes that S.A.T. scores capture a student's academic achievement no more than a student's yearbook photograph captures the full range of her experiences in high school."

To make it perfectly clear that he found the test severely flawed, Judge Walker added, "The probability, absent discriminatory causes, that women would consistently score 60 points less on the S.A.T. than men is nearly zero."

Thomas Kerr

CINTHIA SCHUMAN
Cambridge, Mass., May 8, 1989

The writer heads the National Center for Fair & Open Testing.

BEST COPY AVAILABLE

Subscribe to FairTest Publications

FairTest publications will keep you up-to-date on the growing movement for fair & open testing. You wil be able to do more than just read about FairTest campaigns--you can help make them successful!

_________ *FairTest Examiner*, published quarterly. . . $15.00/year (individuals), $25.00 (institutions).

_________ *Standing Up to The SAT*, by John Weiss, Barbara Beckwith and Robert Schaeffer (192 pages). . . $6.95

_________ *Fallout From the Testing Explosion: How 100 Million Standardized Exams Undermine Equity and Excellence in America's Public Schools*, by Monty Neill and Noe Medina (45 pages). . . $8.95.

_________ *Sex Bias in College Admissions Tests: Why Women Lose Out*, by Phyllis Rosser and FairTest staff, 2nd ed. (40 pages) $7.95

_________ *The Reign of ETS*, by Allan Nairn and Ralph Nader (550 pages). . . $30

_________ *Beyond Standardized Tests: Admissions Alternatives That Work*, by Amy Allina with FairTest staff (15 pages) $5.50

_________ *None of the Above, Behind the Myth of Scholastic Aptitude*, by David Owen (327 pages, lists at $7.95) $5.95

Updates, issued bi-annually (Spring & Fall). $12/year/topic

- _____ Testing Reform Initiatives
- _____ K-12th Grade Testing
- _____ University Admissions Tests
- _____ Teacher/Employment Testing

Order total _____________

Tax-deductible contribution _____________

Total Enclosed _____________

_____ Back Issues of *FairTest Examiner*, $4.00 each: Vol — No. — (All 9 for $32.00)

_____ **Moving.** I have moved/will be moving. My new address is (Please enclose old label):

Name _______________________________

Organization _______________________________ Phone No. _______________

Address _______________________________ ZIP _______________

Please make checks payable to FairTest and mail to:

FairTest, 342 Broadway, Cambridge MA 02139

***Standing Up to the SAT* coming in September from ARCO**

FairTest Book to be Published in Fall

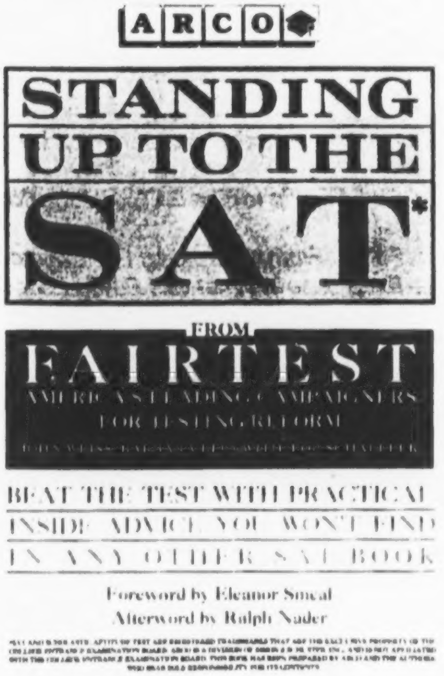
Each year close to two million American teenagers are forced to undergo a painful ritual called the Scholastic Aptitude Test. Now there's a handbook for those who want to use the SAT's flaws to beat, avoid or change the test.

This fall, Simon & Schuster will release FairTest's *Standing Up to the SAT* under the ARCO imprint of its Prentice-Hall division. Designed for high school students, their parents, teachers and advisors, as well as testing reformers, *Standing Up to the SAT* is a compact 192 page paperback priced at $6.95. The book includes a 10 point "Bill of Rights" for reforming the SAT as well as a concise test preparation program. It also describes colleges that have dropped or reduced pre-admission testing requirements. Eleanor Smeal, President of the Fund for a Feminist Majority, wrote the book's foreword, and consumer advocate Ralph Nader penned an action-oriented afterword. Arnie Fege, Executive Director of the National PTA, says the book is "Must reading for students and parents!"

Co-authored by FairTest President John Weiss, journalist Barbara Beckwith and FairTest Public Education Director Bob Schaeffer, *Standing Up to the SAT* will be available as part of a "back-to school special" at Waldenbooks outlets around the nation and in many other stores this fall.

Advance copies are available now to *FairTest Examiner*

(continued on page 13)


ARCO
STANDING
UP TO THE
SAT*
FROM
FAIRTEST
AMERICA'S LEADING CAMPAIGNERS
FOR TESTING REFORM
BEAT THE TEST WITH PRACTICAL
INSIDE ADVICE YOU WON'T FIND
IN ANY OTHER SAT BOOK
Foreword by Eleanor Smeal
Afterword by Ralph Nader

FairTest
National Center for Fair & Open Testing
342 Broadway
Cambridge, MA 02139-1802
(617) 864-4810
(617) 497-2224 (FAX)

ADDRESS CORRECTION REQUESTED

Non-Profit
U.S. Postage
PAID
Cambridge, MA
Permit No. 53278

Please subscribe today.
See page 15.

FairTest
EXAMINER

The National Center for Fair & Open Testing Vol. 3, No. 4 Fall, 1989 ISSN # 0898-2511

New York Report: Ban Mass Testing of Young Children

"NYPIRG's study reveals such a vast misuse, abuse and overuse of standardized tests [through grade two] that there is only one solution to the problem: ***New York State must ban mass use of standardized tests from pre-school through grade two and use alternative and appropriate methods of evaluation.****"*

So concluded the New York Public Interest Research Group after analyzing the results of the first comprehensive survey on test use from preschool to grade two (P-2) in New York State's public schools. The report *Failing Our Children: How Standardized Tests Damage New York's Youngest Students*, found that 75% of school districts use standardized tests on kindergartners, 88% use them on first graders, and 93% of the districts administer them to second graders— despite the absence of a state mandate for so much testing of young children. Nearly one quarter of the districts test students prior to kindergarten.

NYPIRG's findings are based on a survey of New York's 700 school districts, 20% of which responded. The sample was geographically and racially representative of the state.

Dangers of Testing

"This much testing is not only unnecessary, it is harmful for young children," explained NYPIRG's Joan Harris, co-author of the report. "It hurts many boys and girls and damages the curriculum for all children. That's why we have to ban mass standardized testing."

The report, written with the assistance of FairTest, discusses several types of damage caused by the extensive use of standardized exams:

✚ Testing encourages emphasis on a developmentally inappropriate curriculum for young children. This undermines quality education and may turn young children off to schooling altogether.

✚ Individual educational decisions made on the basis of test scores are

(continued on page 6)

ETS ABRIDGES RIGHTS
ETS JUDGE AND JURY
HONESTY IS THE BEST POLICY?
STAND AND DELIVER

Students at Crescenta Valley High School near Los Angeles demonstrating against ETS' refusal to release their AP exam scores following cheating claim. ***Story on Page 11.***

CONTENTS

Shanker Blasts Tests..................2
NC Alternatives 3
Still Above Average 4
News from the States 5
New MMPI Still Invalid............7
10 Myths About the SAT..........8
University Assessment 9
New SAT/ACT Scores
Show Flaws 10
Stand & Deliver II 11
National Merit Scholarships
Called Biased 12
FairTest Book Sales Soar 16

TM017189

FairTest
National Center for
Fair & Open Testing
342 Broadway
Cambridge, MA 02139-1802
(617) 864-4810
(617) 497-2224 (FAX)

***FairTest Examiner*, Volume 3, #4.**
Published November, 1989

Staff: Mike Friedman, Publications Director/Office Manager; Monty Neill, Associate Director; Bob Schaeffer, Public Education Director; Cinthia Schuman, Executive Director; Jamie Souder, Administrative Assistant; Sarah Stockwell, University Tests Coordinator.

Board of Directors: Denise Carty-Bennia, Esq., CUNY Law School, Chair; Dr. Pamela George, North Carolina Central University; Paul Pottinger, Goodsmith and Co.; William Robinson, Esq., Dean, D.C. School of Law; Bob Schaeffer, Treasurer; Steve Solomon, Esq.; Chuck Stone, *Philadelphia Daily News;* Ron Vera, Esq., Barbosa & Vera; John Weiss, Public Policy Initiatives, President.

National Policy Panel: Leonard Beckum, City University of New York; Dr. Beverly Cole, NAACP; Dr. Harold Dent, Association of Black Psychologists; Arnold Fege, National PTA; Dr. Howard Gardner, Harvard University; Dr. Asa Hilliard, Georgia State University; Rep. Matthew Martinez (D-CA); Deborah Meier, Central Park East Secondary School; Sara Melendez, University of Bridgeport; Aurelio Montemayor, Intercultural Development Research Association; Harry Pachon, National Association of Latino Elected and Appointed Officials; Alfredo de los Santos, Maricopa Community Colleges; Eleanor Smeal, Fund for a Feminist Majority; Theodore Shaw, Esq., NAACP Legal Defense Fund; Dr. Gail Thomas, Texas A & M University; Rep. Ted Weiss (D-NY).

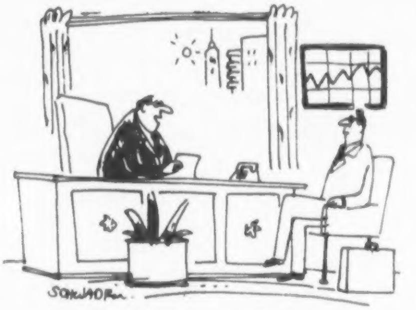
Partial funding for FairTest's work comes from the Golden Rule Corporation, the Rockefeller Family Fund, the Ford Foundation and the Phillip M. Stern Family Fund.

FairTest Provides Testimony to Summit

FairTest was one of a number of organizations invited to present written testimony to President Bush's Education Summit with the nation's governors, which was held in Charlottesville, Virginia on September 27 and 28.

The testimony, entitled *Cautions on Standardized Testing*, is a two page overview of the problems of testing children in grades K-12.

FairTest called on the Summit to endorse authentic methods of evaluation "to ensure that school reform is effective in increasing the quality of education and ensuring equitable access to improved education." For a copy of the testimony, send an SASE to Summit Testimony, FairTest, 342 Broadway, Cambridge, MA 02139.



"After 20 years of schooling, your aptitude test shows that you're skilled at just one thing—taking tests."

From *Phi Delta Kappan*

Upcoming Speaking Engagements

FairTest staff members can be heard speaking at the following events in the next several months:

Dec. 11: Cincinnati Public Schools second annual education convention.

Feb. 10: Keynote address to the annual meeting of the Hawaii Association for Supervision and Curriculum Development.

Feb. 24: American Association of School Administrators annual convention in San Francisco, a talk titled "Reducing Standardized Testing and Developing Alternatives."

Quote of the Month

"I would call for an immediate end to standardized tests as they are now. . The tests are not relevant to the outside world . . .Teachers are forced to spend a lot of time helping students to do better on standardized tests, which are really narrow and determine curriculum . . . To test for depth of knowledge, you have to get away from things that are machine gradeable."

—**Albert Shanker, President of the American Federation of Teachers**, speaking at the Educational Testing Service's 1989 Invitational Conference, held in New York City, October 28, 1989.

Grade School Testing

Alternatives Implemented in North Carolina

This year, North Carolina has begun to implement alternatives to standardized achievement tests. These alternatives were mandated by the 1988 state law that banned achievement tests in grades one and two (*Examiner* Vol. 2, No. 2). North Carolina's experience demonstrates that authentic assessments can be developed and used on a statewide basis.

According to Norma Kimzey the new assessments, based on the state's developmentally appropriate curriculum, were designed and field-tested in 1988-89. They include observational check-lists on which teachers note each student's progress toward developing knowledge and skills in reading, writing, oral language and math. All teachers have been given detailed descriptions of what and how to assess. Kimzey, an educational specialist for the North Carolina Department of Public Instruction, helped write the *Communications Skills* portion of the assessments (the other part is *Mathematics*) and is helping train teachers in its use and implementation.

The new assessments include instructions on such tasks as compiling a portfolio of student work. Because this approach is radically different from both standardized testing and conventional teaching methods used by many teachers, Kimzey explained, "we are showing teachers how the new process links instruction with assessment and is based on the child's own exploration of the world and construction of knowledge."

Training and Implementation

This past summer, a group of "master teachers," selected on the basis of prior involvement in using the state's new curriculum and interest in the new assessments, were trained for one week in how to use the new procedures.

> *"We are showing teachers how the new process links instruction with assessment."*
>
> *--Norma Kimzey*

Each master teacher in turn worked with a group of 25-30 teachers in a week-long program designed to prepare them to become lead teachers in their schools. The lead teachers then direct 12 hours of in-service training for the remaining teachers.

The Department has produced videos showing the assessments in use in classrooms. Specialists are also meeting with education faculty in the state's colleges and universities so that the new procedures can be incorporated into the training of new teachers.

Problems in Implementation

The implementation has uncovered some problems. The most serious, according to Kimzey, is that teachers have to learn to assess as they teach. Some teachers are not comfortable exercising their own judgement, which the assessment procedures require. Thus, the new assessment has become a major tool of teacher re-education and of implementing the developmentally appropriate curriculum.

Resource limitations have been another problem. In each of the state's eight regions, there are two specialists, one each in communications and math. The specialists provide assistance and training in all educational areas from grades K through 12. The small number of staff, the wide range

(continued on page 6)

GLASBERGEN

"My guidance counselor says that I have the highest I.Q. in my school. She says I.Q. stands for 'irritation quotient.'"

From *Phi Delta Kappan*

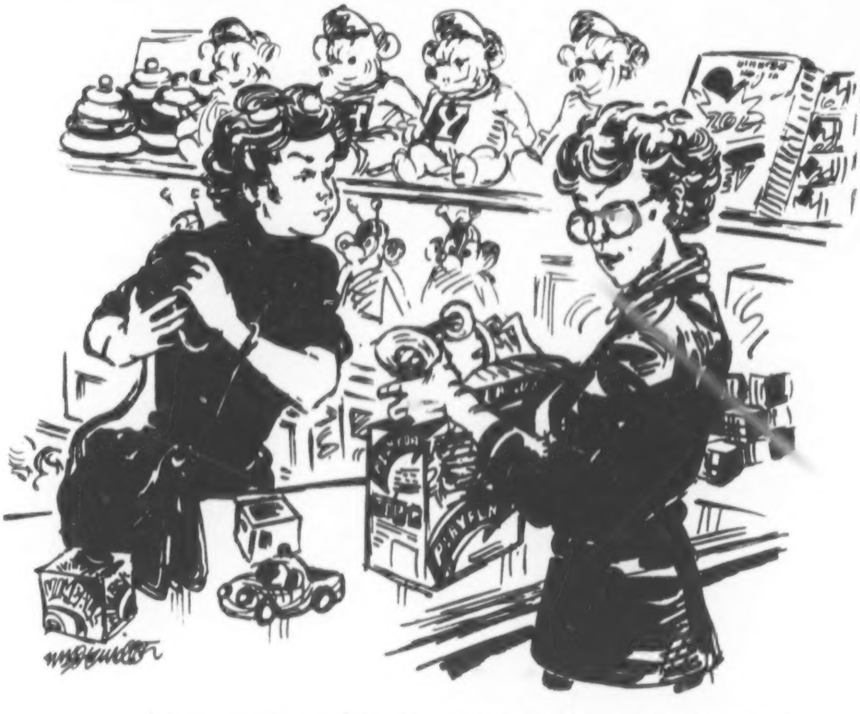
Grade School Testing

Science Assessment

Connecticut's Education Department has received $1 million from the National Science Foundation to develop performance-based student assessments in high school science. Teachers from seven states are being trained in the new method.

"We believe the time has come to develop assessments that are catalysts for the kind of learning that we value," wrote the department in a report on the assessments. "The model currently in place—atomistic tasks, passive learning, and primarily convergent thinking—has been too well served for too many years by multiple-choice testing" (*Education Week*, Sept. 13, 1989).

The new methods will measure performance on a series of tasks over an extended period of time. Connecticut officials hope that the new assessments will help reform the way science is taught, as well as largely replace multiple-choice tests. The state plans to implement the new methods statewide in 1991, but officials from other states whose teachers were being trained said they would wait and see before adopting them.



"Two months with this and they blow their preschool entrance exams right out of the water."

From *The New Yorker*

Still Above Average

In a new report, Dr. John Jacob Cannell of Friends for Education has found that 48 of the 50 states and 80% of the school districts in the country still report test scores above the national average. Cannell raised a furor in 1988 when he first reported such results, which were largely confirmed by a more comprehensive study conducted by the Center for Research, Evaluation, Standards and Student Testing (CRESST) at UCLA.

In his new study *The Lake Wobegon Report: How Public Educators Cheat on Achievement Tests*, Cannell charges that the high scores exist because educators see test items and teach them to their pupils. CRESST had noted that teaching to the test was a major cause of the "Lake Wobegon" effect (where "all the children are above average").

Though Cannell's solution to the problem is increased test security, CRESST director Eva Baker called for replacing standardized tests.

Cannell has again called on the Federal Trade Commission and the 50 state attorneys general to investigate the test companies for "unfair and deceptive trade practices." In addition, he asked Education Secretary Lauro Cavazos to issue guidelines so that testing will not "deceive the American public." Cannell filed similar consumer complaints in the 50 states in 1988, but no action was taken.

The Lake Wobegon Report *is available for $15.00 from Friends for Education, 600 Girard Ave., NW, Albuquerque, NM, 87106.*

BEST COPY AVAILABLE

Grade School Testing

Testing in the States

Arizona Superintendent Calls for Alternatives

Tests of young children "rarely yield valid and meaningful results," concluded a report on early childhood education issued by Arizona School Superintendent C. Diane Bishop. "Readiness testing," she added, inappropriately excludes children. Moreover, teachers "waste valuable time 'teaching to the test'" which "has a serious, detrimental effect on the curriculum and the children." The superintendent called for "the development of alternatives to standardized tests for young children," and a developmentally appropriate curriculum in line with current knowledge about how young children learn.

Meanwhile, Arizona's Education Department has introduced legislation to supplement the state's mandated testing program with "performance-based" and other alternative assessments. The state currently administers a multiple-choice test to every student in grades two through eleven. This testing will continue, but they claim the amount will be reduced by two-thirds. Tests in grades one and twelve were dropped a year ago (see *Examiner* Vol. 2, No. 3).

Lois Easton of the Education Department told *Education Week* (Oct. 11, 1989) that the new assessments would provide more information about student performance and encourage districts to emphasize higher-level abilities in their curricula. The standardized exams, she said, only cover one quarter of the curriculum that they are supposed to test.

Illinois Passes Test Limits

The Illinois legislature has passed a bill that will limit to five hours per student per year the amount of time devoted to state mandated standardized testing. Each district will be required to report annually "the amount of time spent per pupil on the state student assessment testing and the total amount of time spent per pupil on the local student assessment testing." Gov. James R. Thompson used an amendatory veto to change other parts of the bill, but the testing limit was not affected.

Tennessee Reduces Testing

To reduce the number of tests and time spent on them, Tennessee has adopted a customized version of the Comprehensive Test of Basic Skills that combines norm-referenced with criterion-referenced items. The test is optional in grade one, mandated in grades two through eight and ten. If a student scores high enough on the CTBS in grade eight, the state requirement that students must pass a proficiency exam to graduate can be waived. The state also believes that providing multiple forms of the combined test will result in more broadly defined subject areas, making teaching to the test and narrowing of the curriculum more difficult. The format remains multiple-choice.

Kentucky Adopts CTBS

Last year, Kentucky dropped its criterion-referenced Kentucky Essential Skills Test, which had been given in all grades. However, it adopted the norm-referenced Comprehensive Test of Basic Skills (CTBS), which was given last spring in grades three, five, seven and ten. The results showed Kentucky to be one of the 48 states that are "above average."

New HS Comp. Test in Okla

Legislation just passed in Oklahoma will expand state testing, and initiate a high school graduation competency exam beginning in 1993.

Education or Coaching?

Riverside Publishing, maker of the Iowa Test of Basic Skills, now offers "Improving Test-Taking Skills." In two sets of kits, one for grades four through six, the other for grades six through nine, Riverside offers a "course" in how to take tests. Not content with informing students about "following test directions" or "using (test) time wisely," the kits essentially instruct students to reduce their abilities and knowledge to a format conducive to test-taking. They suggest that "being tested is similar to being taught" and provide "structured exercises for strategic skills practice." A school kit, containing "30 booklets for each of 15 skills; 1 Teachers Guide" and answer forms, costs $488.70 each for grades four through six or six through nine. Order now so that you, too, can turn education into test coaching.

Grade School Testing

Report Calls for P-2 Testing Ban in New York

(from page 1)

frequently wrong because tests of young children are often unreliable and invalid.

✚ Many children are tracked on the basis of test scores and remain in these tracks for their entire school careers. Low-income and minority group children are disproportionately placed in low-level tracks in which they fall continually farther behind their peers.

"School districts are also misusing the exams by using tests for purposes for which they have not been validated," observed Blair Horner, NYPIRG's Legislative Director.

For example, the Gesell School Readiness Test, which is supposed to gauge "readiness" for kindergarten or first grade, is often misused as a disabilities screening instrument. The Gesell, however, has not been validated by its maker as a "readiness" exam, let alone as a "screening" test.

"The continued use of IQ tests in New York is particularly outrageous," said Susan Harmon of the American Reading Council. "These tests have been banned in California for use on black children because they are racially biased and lack educational validity, but they are still being used throughout New York." The survey reported, for example, that one quarter of the schools which tested second graders used I.Q. tests, often to track students or to place them into special programs.

> *"The state does not require the districts to administer all these tests. They do it on their own."*
>
> *--Joe Sammons, NYPIRG*

Testing Not Required

"We were very surprised to find that districts were doing so much testing," said NYPIRG's Joe Sammons, who analyzed the survey results. "The state does not require the districts to administer all these tests. They do it on their own."

The report noted that the only testing of young children mandated by New York State law or regulation is the Linguistic Abilities Battery (LAB), which is given for placement into bilingual education. The state does require all children to be screened for placement into special education or "gifted and talented" programs, but that screening does not have to be in the form of a test. Under federal Chapter I guidelines, all students in Ch. I remedial programs in grade two and above have to take a standardized test, but that is the only Ch. I testing requirement.

Winning a Ban

NYPIRG, FairTest and the American Reading Council have already convened several meetings in New York to discuss forming a coalition that would work toward banning the mass use of pre-kindergarten through grade two standardized testing in New York. The next meeting will be on December 8. For more information on the coalition, call Joan Harris at NYPIRG (212) 349-6460 or Susan Harman at ARC (212) 619-6044. Copies of *Failing New York's Children: How Standardized Tests Damage Our Youngest Students* are available for $5.00 from NYPIRG, 9 Murray St., NY, NY 10007.

New NC Alternative Assessments

(from page 3)

of their responsibilities, and their diverse backgrounds slow the amount of training that can be done. Moreover, with a budget of only about $150,000 for the state, they lack the resources to provide more training and materials.

Another problem is that, under federal law, all states must administer a norm-referenced test to all second graders who are in Chapter I remedial programs. In North Carolina, however, the new assessments and teacher recommendations together will be the basis for placing income-eligible children in the remedial program.

Finally, while assessments based on children's own activities should not be biased by race or class, some teachers inevitably will be prejudiced. This issue, said Kimzey, needs to be addressed. What will be needed, she thought, is some form of monitoring that could analyze assessment results by demographic categories such as race or income.

An expanded version of this article is available. Send an SASE addressed to N.C. Alternatives, FairTest, 342 Broadway, Cambridge, MA 02139.

Employment Testing

New MMPI—Still No Validity for Employment

In response to complaints that many test items were outdated, sexist and pre-occupied with religion and personal bodily functions, National Computer Systems (NCS) of Minneapolis has introduced a new version of its venerable Minnesota Multiphasic Personality Inventory (MMPI). This new edition will replace the original test, first developed at the University of Minnesota in the late 1930's.

Over 250 of the items have been changed from the original version, and attempts have been made to make the test gender-neutral. But by undertaking such a major revision, the makers of the MMPI are effectively admitting that an exam in wide use since 1943 was sexist and invalid for much of the population of the United States.

The new release, MMPI-2, consists of 567 statements, to which the test-taker must reply "true" or "false." The items range from "I prefer a bath to a shower," and "Sometimes I tease animals," to "I never tell a lie." By evaluating responses to the statements, the makers of the MMPI-2 allege that it can detect liars, people with substance abuse problems and mentally unstable individuals. Originally intended only for psychological evaluation, the MMPI-2 is now widely used for job screening as well, especially for "high stress" jobs such as fire fighters or airline pilots.

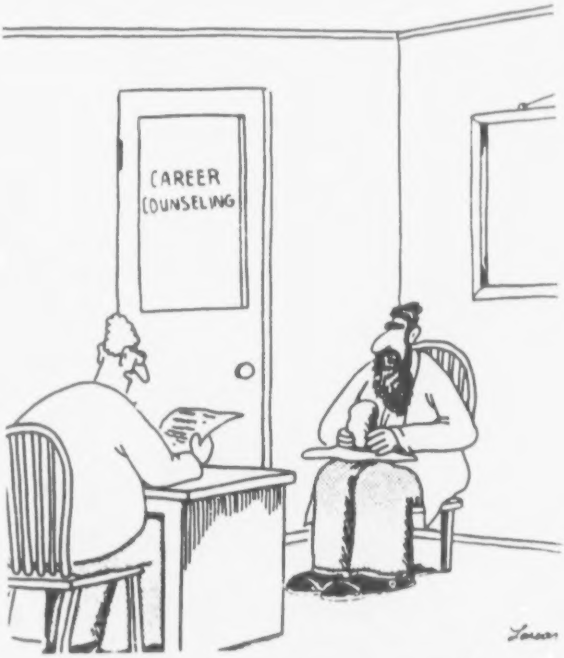
The original MMPI was normed on a group of 700, mostly rural, Minnesotans, with an average age of 35 and an average 8th grade educational level. Most of the men were skilled or semi-skilled workers while most of the women worked in the home. Although the new normative group consisted of a diverse sample of 2600 people from across the country, the MMPI-2 still has not been validated for use as an employment screening tool. Kevin Anderson of NCS told FairTest that the MMPI-2 would be more valid than the old instrument for use in making employment decisions, though he admitted that no validity studies have been done to support this use of the test.

> *". . . I think it's a mystery to the people who developed it why it works."*
>
> *--Psychologist Val Arnold, on the MMPI-2, speaking to the Wall Street Journal.*

Validation of a test demonstrates how accurately or meaningfully it measures what it claims to measure. "Construct validity" indicates how well the test corresponds to an underlying, theoretically defined trait. Yet independent consulting psychologist Val Arnold of Minneapolis told the *Wall Street Journal*, ". . . I think it's a mystery to the people who developed it why it works."

In a review of the original MMPI in the *9th Mental Measurements Yearbook*, reviewer Dr. David Nichols criticized use of the MMPI for making employment decisions. Nichols wrote "There are substantial reasons to

(Continued on page 13)



"Well, Mr. Cody, according to our questionnaire, you would probably excel in sales, advertising, slaughtering a few thousand buffalo, or market research."

Gary Larson, *The Far Side*

University Testing

Ten Myths About the SAT

Though the public has become much more savvy and aware of testing issues in recent years, misconceptions about the SAT abound. The following are ten common myths about the SAT, heard in reaction to FairTest's new book, *Standing Up to the SAT*.

1) The SAT gives all students an equal shot at college admission.

Because it's not a level playing field, the SAT denies blacks, Hispanics, Native Americans and women equal opportunities for higher education. Research shows that when admissions offices place heavy emphasis on SAT scores — as many do — the number of blacks and low-income students admitted goes down. What's more, using scores to award scholarships prevents minorities and women from getting their fair share of badly-needed tuition aid. Through its culturally-specific questions and format, the SAT compounds discrimination against disadvantaged groups.

2) The SAT is only the messenger: score gaps merely reflect differences in students' academic backgrounds.

Even when matched for academic preparation, there are still large gender and race gaps in SAT scores. That's one reason why a federal court, after reviewing the test-makers' best arguments, which removed the effect of variables such as ethnicity, parental education, high school classes and proposed college major, found that "...[U]nder the most conservative studies presented in evidence, even after removing the effect of these factors, at least a 30 point combined differential [between males and females, out of approximately 60 points] remains unexplained."

3) SAT scores are precise, like inches on a yardstick.

The College Board's *ATP Guide* says "Users learn to understand and appreciate the meaning of a score of 430 in the same way that they have learned . . . . the meaning of . . . . 14 inches. . ." But it notes elsewhere that the SAT's margin of error is *65 points*. Even more incredibly, the test-makers admit that two students' scores must differ by at least *138 points* before anyone can be sure that their abilities differ. If the SAT is a "common yardstick," as the test-makers say, it must be made of elastic.

4) At least the SAT is more accurate than high school grades.

On the contrary, despite all the differences between courses and grading standards, high school grade point average is still the best single predictor of first year college grades—which is all the SAT claims to predict. This shows just how inaccurate the SAT really is.

5) But colleges still need test scores to make admissions decisions.

Clearly the 30% of schools which admit virtually all of their applicants don't need any scores at all. More competitive colleges can follow the example of Bowdoin which has found that the diversity and quality of its students improved after it made the SAT optional 20 years ago. Admissions officers at the fewer than 50 schools that reject more than half their applicants have many other ways to deal with differences in high school curriculum and quality.

6) When girls and boys are matched by ability, the score gap on the SAT-Verbal section disappears.

The College Board's way of "matching by ability" is to match test-takers' SAT scores! So if, as growing evidence indicates, the test is flawed, then matching by scores is also skewed. Even ETS admits this argument is "circular." The real cause of the gender gap is bias in the test: girls take substantially more English in high school and college and earn much better grades, but still receive lower Verbal scores. If the test were measuring accurately, more girls would score high on the SAT-Verbal section than boys, but the opposite is true: only 45% of those who score 600 or more are girls.

7) Test coaching doesn't work.

Studies show that good coaching programs can raise a student's scores by 100 points or more. Many of these courses are very expensive ($500 and up), and teach little more than "test-wiseness" strategies specific to the SAT. The fact that short-term coaching works undermines the test-makers' credibility because they claim that the SAT measures skills and "aptitudes" learned over a long period of time. It also adds another income-related bias to the test: a black woman from a low-income family is triply discriminated against because of her race, gender and inability to pay for coaching. Moreover, why do both ETS and the College Board sell test preparation products if coaching doesn't work?

University Testing

Ten Myths About the SAT

8) The test-makers' "exhaustive bias reduction procedures" guarantee that the test is fair.

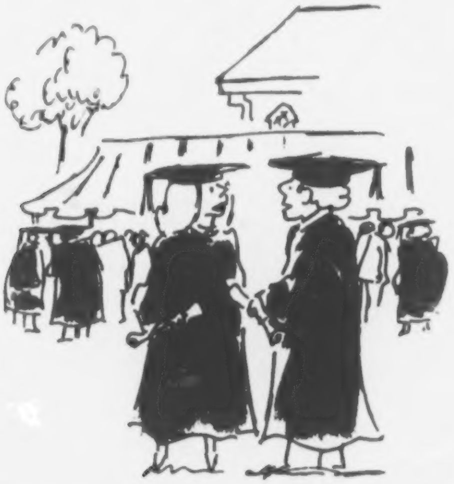
Whatever else their procedures may be doing, they are not eliminating biased questions from the test. Words specific to the activities of well-to-do Americans such as dividends, stockholders, royalties and chess continue to appear on the test. Researcher Phyllis Rosser found 40 questions on the June 1986 and November 1987 SAT that had extreme bias in favor of one gender or the other: over 10% of the questions on these two tests combined favored boys. The claim that bias has been eliminated from the test is little more than a public relations smoke screen.

9) The SAT measures what you need to know in college.

The SAT is a mind game that has nothing to do with skills necessary for higher education: it tests a tiny range of techniques, mainly how quickly you can choose among four or five answers without thinking deeply about any of them. Some of the many qualities you need in college that the SAT cannot measure are writing ability, strategic reasoning, higher order thinking skills, experience, persistence and creativity.

10) The SAT sums up how smart you are.

Despite its name, the Scholastic *Aptitude* Test is not a test of aptitude. Intelligence is a far broader, and more elusive quality than any multiple choice test can measure. The SAT is little more than a test of how many math tricks and unusual words you've learned: it doesn't allow for subtleties. This may be one reason that girls average lowerscores on the test: research shows that girls tend to think more complexly than boys, a quality that hurts them on the SAT.



"Thank heaven we didn't have to take SAT's to get out of here."

From *The Chronicle of Higher Education*

Challenges in Texas and Florida

University Assessment

Increasingly, community colleges, some four year schools, and a number of state college systems have come to rely on standardized tests to measure the "basic skills" of their students, particularly those attempting to move up to Junior-year classes. Current controversies over cut-off scores on skills tests in Texas and Florida illustrate some of the problems caused by reliance on these multiple-choice instruments.

In Texas, up to 93% of black and 87% of Hispanic students are expected to fail the Texas Academic Skills Program (TASP) test, which every student must pass before enrolling in upper-level classes. The Mexican American Legal Defense and Education Fund (MALDEF) is considering a lawsuit against the state because of concerns that the test will limit access to higher education for Hispanic students. When the test was first proposed, some supporters argued that it would identify students who could be helped through remediation. But MALDEF pointed out the likelihood that testing would be implemented but adequate funding for remediation would not be provided. That is exactly what happened.

As an example, MALDEF attorney Al Kauffman says that the typical large community college in Texas is likely to receive about $250,000 per year for remedial programs. After opening an office and hiring an administrator, there will only be

(Continued on page 13)

University Testing

Annual SAT and ACT Scores Show Continued Flaws

This year's admissions test scores for entering college students show basically no changes from last year. Though the SAT's manufacturers attempted to deflect attention by focusing on long-term trends, the gap between males and females and between the richest and poorest SAT test-takers actually increased slightly from 1988 to 1989. The widening gender and income gaps and the persistence of racial and ethnic gaps demonstrate that the Educational Testing Service, College Board and American College Testing Program have failed to address bias in their products.

This year, in an effort to quiet criticism, a College Board spokesman explained to a reporter from the *Arizona Republic* how ETS bias reduction techniques weed out SAT items that are easier for some groups than others. He explained to a reporter that an analogy featuring the word "emancipation" was eliminated after blacks registered a higher success rate on the item than expected. Yet questions using terms such as "dividends" and "stockholders" regularly end up on the test, despite the same bias reduction techniques.

1989 SAT AVERAGE SCORES BY GENDER (400-1600 scale)

	Verbal	Math	Combined
Male	434	500	934
Female	421	454	875

1989 SAT AVERAGE COMBINED SCORES BY ETHNICITY AND INCOME (400-1600 scale)

White	937
Asian Americans	934
Latin Americans	825
Native Americans	812
Mexican Americans	811
Puerto Ricans	766
Blacks	737

over $70,000 - 996
$ 60,000-70,000 - 959
$ 50,000-60,000 - 944
$ 40,000-50,000 - 924
$ 30,000-40,000 - 898
$ 20,000-30,000 - 870
$ 10,000-20,000 - 829
below $10,000 - 780

(from *1989 National Ethnic/Sex Profiles, College-Bound Seniors*, New York: The College Board, 1989)

1989 ACT AVERAGE COMPOSITE SCORES BY GENDER (1-36 scale)

Male	19.3
Female	18.0

1989 ACT AVERAGE COMPOSITE SCORES BY RACE/ETHNICITY AND INCOME (1-36 scale)

Asian	19.9
White	19.4
Puerto Rican	16.6
Mexican American	15.4
American Indian	14.7
Black	13.6

over $60,000	20.7
$50,000-59,999	20.1
$42,000-49,999	19.7
$36,000-41,999	19.2
$30,000-35,999	18.8
$24,000-29,999	18.4
$18,000-23,999	17.8
$12,000-17,999	17.2
$ 6,000-11,999	16.0
less than $6,000	14.2

(From *The High School Profile Report, Normative Data, ACT High School Profile Report, HS Graduating Class 1989*. Iowa City, IA: American College Testing Program, 1989)

University Testing

Calif. Students Sue ETS—"Stand and Deliver II"

It's "Stand and Deliver" all over again for students at Crescenta Valley High School in Southern California. That 1988 movie, starring Edward James Olmos, depicted the struggle of Mexican-American students wrongly accused of cheating on an Advanced Placement Calculus exam. In a new case that shows the test-makers' continued disregard for students' rights, the Educational Testing Service (ETS) has again invalidated the scores of honest test-takers.

7 Cheaters, 72 Victims

The problem began last May when 79 students in La Crescenta, California took Advanced Placement exams in eight subjects. After test-takers reported improprieties during the Biology, U.S. History, and English Language and Composition tests, seven students came forward and admitted cheating. Further investigation by the high school revealed no evidence of wrongdoing by anyone other than the seven students. In a letter, school officials assured the 72 other students that their exams would be scored and validated by ETS "without any cloud whatsoever."

Approximately six weeks later, when test scores should have been released, ETS announced that it was conducting a further investigation. After a brief inquiry, the same ETS employee who investigated the "Stand and Deliver" case assured anxious students that their scores would soon be released. They never were.

> *". . .there's no fair system for resolving serious disgreements with ETS."*
>
> *--Sheryl Wirt, parent leader at Crescenta Valley H.S.*

ETS Reverses Itself

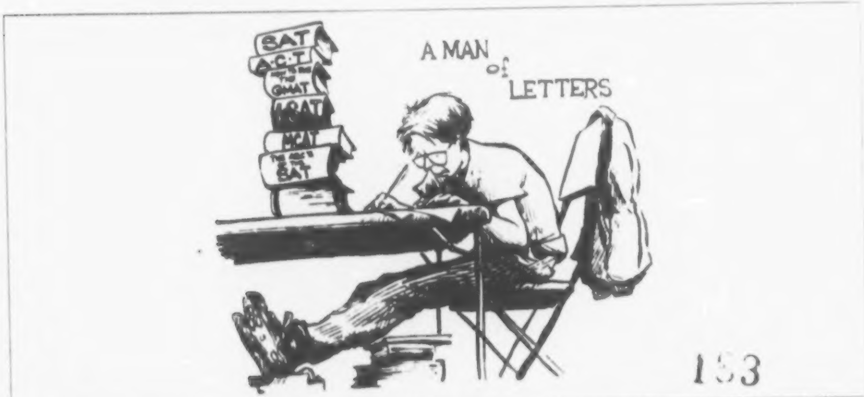
On August 3, after two and a half months of waiting for test results, students were informed by letter that ETS was invalidating their scores due to "administrative irregularities." To the dismay of many students, particularly those who were leaving for college in August, ETS mandated retests. Faced with the choice of studying for the grueling exams over again or standing up for their rights, many students boycotted the retests. With the help of their parents, twenty seven students sought a temporary restraining order blocking the retests. A federal district court judge denied the request, but scheduled a trial for October 31.

No Due Process

In their suit to reinstate the original scores, parents and students maintain that ETS violated California laws designed to provide due process protections to test-takers — the very laws enacted following the "Stand and Deliver" debacle. These protections require the test-maker to provide students with timely notice of any delay in the reporting of test scores, a complete summary of all information relied upon to withhold test scores, and an opportunity to respond to the test-maker's charges prior to a final decision. Parents and students say these procedures were never followed.

Lesson In Power

For Sheryl Wirt, one of the Crescenta Valley parent leaders, dealing with ETS has been a lesson in the abuse of power. "We were shocked to discover that there's no fair system for resolving serious disagreements with ETS. They've been completely unresponsive and arbitrary." Regardless of the outcome of their lawsuit, the group has vowed to press for changes in the balance of power between test-takers and test-makers, including independent monitoring of the testing industry. FairTest will continue to work with this group to build support for the "Test-Taker's Bill of Rights," as outlined in *Standing Up to the SAT*.



EDUCATION DAILY

The education community's independent daily news service

Vol. 22, No. 206 · Wednesday, October 25, 1989

Watchdog Group Questions National Merit Exam's Fairness to Girls

A standardized exam used to screen semifinalists for National Merit Scholarships is "denying girls their fair share" of awards, charges FairTest, the National Center for Fair and Open Testing.

In a state-by-state analysis of this year's scholarship semifinalists, FairTest found that by a nearly 2-to-1 ratio, boys outnumber girls competing for the scholarships totaling $23 million, despite girls' higher grades in both high school and college.

Of the 15,467 National Merit semifinalists announced this fall, 58 percent are male and 36 percent are female, FairTest found. The gender of the remaining students could not be determined by their names.

FairTest released its annual survey late last week, just before high school juniors across the country took the Preliminary Scholastic Aptitude Test/National Merit Scholarship Qualifying Test (PSAT/NMSQT), the sole factor used to select semifinalists.

Scores Not Only Factor

Marianne Roderick of National Merit Corp. said there are no easy answers to the issues raised by FairTest, but she noted that scholarships are not awarded solely on test performance.

"We use the PSAT only as a screening device at the initial level," Roderick said. "Our semifinalists have to demonstrate much more academically in the long run." Other criteria include academic achievement, extracurricular activities, recommendations from principals and student essays.

Boys may receive more scholarships, she said, because they take more difficult classes than girls, especially in science and mathematics. "We think that is changing as girls are encouraged to take more of such courses, but changes often take time."

A spokesman for the Educational Testing Service, which designs the test, noted that every effort is made to insure questions are unbiased. "We analyze every question according to gender and ethnic background and if we see a gap, we carefully pick apart the question, to find out why," said ETS spokesman Tom Ewing. If the question remains troublesome, he said, it is discarded.

"We take very seriously trying to make the test as fair as possible," said Ernest Kimmel, ETS executive director of test development.

Gender Gap Persists

FairTest, however, says ETS safeguards aren't working. "The questions statistically favor males," said FairTest Executive Director Cinthia Schuman. "Even when girls and boys are matched by the courses they take, we still find a gender gap. But all other measures of academic excellence show females outperform males."

As an example, Schuman pointed to a five-year survey of high schools in the Chicago area, where every year between 53 percent and 60 percent of valedictorians were female. "By continuing to use a biased and inaccurate test, National Merit is denying girls their fair share of scholarships," she said.

FairTest's report is just its latest volley against the scholarship program, charging that the standardized tests discriminate against girls (ED, May 11). ETS has defended its tests against previous criticisms, saying girls score lower than boys for complex reasons (ED, April 28).

FairTest said a New York ruling backs its position that relying on Scholastic Aptitude Test (SAT) as a sole determiner of scholarship awards is discriminatory. In *Sharif v. New York Education Department* (88 CIV 8435), U.S. District Court Judge John Walker issued a preliminary injunction barring the state from awarding scholarships based exclusively on SAT scores (ED, Feb. 7).

--David Schumacher

(Re-typeset from Education Daily, October 25, 1989)

Multiple-Choice Assessments Harm Minorities *(from page 9)*

enough money left for a few large classes focusing on test preparation drills. Funds are simply insufficient to provide any of the kinds of remediation that have proven effective, such as smaller classes and individual attention.

Last summer, students in Texas took advantage of a loophole in the law which allowed them to take three credit hours of college-level work over the summer rather than pass the TASP test. Community and junior colleges in the state saw record enrollments as students rushed to avoid the test: Collin County Community College, for example, reported a 79 percent increase in enrollment.

In Florida, the NAACP has threatened to sue over a scheduled hike in the cut-score for the College Level Academic Skills Test (CLAST), which, like the TASP test in Texas, must be passed by all public college students before enrolling in upper-level college coursework. The new cut-score is touted by proponents as a way to raise academic standards. However, Robert H. McCabe, President of Miami-Dade Community College, said in the *Chronicle of Higher Education* that the higher cut off will have a "devastating impact" on minority students. An estimated 79% of black and 70% of Hispanic students will fail. Raising the cut-off score will have additional negative educational effects. McCabe says that the current cut-score has already prompted students to narrow their class choices from those of general educational interest to those that will help prepare them for the test.

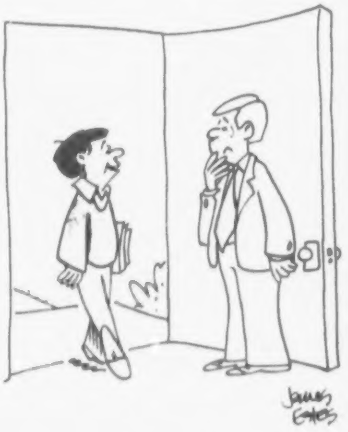
Avoiding tests is not limited to students: according to the *Chronicle of Higher Education* (10/4/89), a Florida legislator found that his fellow lawmakers weren't interested in sampling the CLAST to get "a better idea of reality." Rep. Tom Tobiassen, a supporter of the test, said, "A lot of them kind of balked" when he suggested that they take the test to "see what it's like" and to get some perspective. Tobiassen thinks the present form of the CLAST might be "a little too tough."

New MMPI *(from page 9)*

question the suitability of the MMPI in personnel selection," and "use of the MMPI in selection for non-critical occupations would seem to invite credulous and immoderate use."

Despite the absence of validity studies, Arnold told the *Journal* he thought the new version was an improvement. After observing that the old MMPI "overidentified conservative Christians as having mental problems," Arnold explained "If you believe in God and you believe in miracles, there's nothing particularly crazy about that. But if it's not part of your religious faith, you might be suffering delusions."

There is no way of knowing how many employers are using either version of the MMPI as a sole criterion for job screening, in clear violation of the test publisher's guidelines. Obviously a test which experts find "mysterious" and which claims to be able to spot unstable people by how they prefer to clean themselves, is ripe for serious abuse. The MMPI is a prime example of how a badly-normed, unvalidated instrument can become widely used to make important decisions about people's lives.



"I finally got 100 on a test, Dad . . the SAT."

From *Phi Delta Kappan*

THE BIRMINGHAM NEWS
Sunday, May 28, 1989

Standardized tests are taking over the direction of schools

"If it moves, test it!"

This has become the growing mentality infiltrating our schools during the past decade. The amount of testing of American school children is greater now than at any other time in history.

Introduced in the 1920s, standardized tests have grown to dominate our educational systems. Once used as one of several alternative methods to assess children's achievement, competency and basic skills, normed standardized multiple-choice tests now are the primary criteria used by most schools for making decisions that affect not only a child's future but also the development of curriculum and the quality of teaching and the school environment.

Curriculum dominated by testing industry

Education Secretary Lauro Cavazos has recently declared that the education reform movement has lost its momentum and children's performance has stagnated. More money is being spent but academic progress is at a standstill, and the high school dropout rate continues to climb. But, Cavazos doesn't have an answer to what has happened.

The answer to what has happened may, though, be quite evident. Using standardized test scores as a primary criterion for making important educational decisions has greatly weakened our education system. The tunnel-visioned reliance on the standardized test as the primary or sole assessment tool to substantiate child, teacher, school competency and educational worth has given us a curriculum now dominated by the unregulated testing industry.

Each year we eagerly await the local newspaper's publication of test score results. The teacher's and the school's worth are on the line. One line of print seems to be the justification for judging a year's work. In their efforts to get higher test scores, school systems have developed a paranoia that is severely constricting their curricula. We are at the crossroads for making important decisions regarding curriculum, assessment and the school climate in which our children are expected to adapt, learn and develop.

During the 1986-87 school year, 105 million standardized tests were given to 39.8 million school children. According to a study by the National Center for Fair and Open Testing (FairTest) in Cambridge, Mass., that's an average of more than two and one-half standardized tests per pupil each year. By the time a child graduates from high school, he has the likelihood of having been subjected to 30 such tests. And it starts in pre-kindergarten and kindergarten, where approximately 2 million youngsters are now put through formal screening each year. The No. 2 pencil has been added to the crayon box.

PERSPECTIVE: A CAMPUS VIEW

Tom Jambor

As the quantity of testing has increased, so has the number and the intensity of complaints. Let's take a look at some of the recurring criticisms and findings accumulated by FairTest:

● Standardized tests often produce results that are inaccurate, inconsistent and biased against minorities, females and students from low-income families.

Tests that correlate readiness with later school performance are unreliable. Some 30 to 50 percent of the children who are identified as not ready to start a grade will have been placed in that category erroneously.

Students whose test scores are low are presumed to be unable to master complex material and are often subjected to a watered-down curriculum.

Because of their scores on unreliable exams, a large number of racial and ethnic minorities, as well as a growing number of the majority population, are placed into programs for "slow-learners" under the assumption that they cannot or will not learn. This "tracking" creates a self-fulfilling prophecy.

● Standardized tests focus time, energy and attention on the simpler skills that are easily tested and away from higher-order thinking skills and creative endeavors.

They emphasize isolated learning and assume a single "right" answer, rather than allowing for complex answers or multiple ways of arriving at them.

Tests are timed, thus placing value on thinking quickly rather than on thinking profoundly. Quick recognition of isolated facts is emphasized and measured, without regard for the student's ability to synthesize information, solve problems, think independently or generate ideas.

High-achieving students become frustrated and bored by a restricted, narrow curriculym that has been watered-down in response to standardized tests; an increase in drop-outs often follows.

● Standardized tests tend to corrupt the process of teaching and learning, often reducing teaching to mere preparation for testing.

The Test has become the curriculum. Time is taken away from real teaching when the curriculum is narrowed simply because time is needed to prepare students for testing and for taking the test. Teachers, pressured to raise test scores, have succumbed to excessive use of drill or isolated skills.

As states skew their curriculum to meet the demands of testing programs, parts of subject areas that are important tend to be ignored.

We must remember children go to school not just to learn basic academic skills, but also to develop the personal, intellectual and social skills that will enable them to become happy, productive and responsible members of our society. However, important acquisitions such as critical thinking, reasoning, social competence, self-esteem, desire to learn, and self-discipline are not easily measured by standardized tests, and therefore are virtually ignored or given minor importance in schools with mandated testing programs.

Our fascination and obsession with normed standardized tests is putting us on a serious collision course with the future of education and children.

When we judge achievement only in terms of outcomes of standardized tests, we create a classroom climate that is profoundly alienating to both children and teachers.

Should relegate tests to subordinate role

Normed referenced testing, if used at all, must be relegated to a subordinate role in the total assessment picture. Assessment alternatives are an absolute necessity if we are to improve instruction and learning. With a wide range of assessment opportunities, teachers will be allowed to pinpoint not only what a given student knows, but how that student learns best.

Samples of student work over an extended period of time, teacher-made and criterion-referenced tests, accumulated grades on daily activities, teacher observations of behaviors, actions and performances, and the teachers' professional judgment are a few alternative sources for gaining a better total understanding of the development of the whole child.

The curriculum must be put back into the hands of the professionals who have been trained to work with children and, given the opportunity, construct appropriate programs and assessments that account for individual differences and all children.

Tom Jambor is associate professor-early childhood development in the University of Alabama at Birmingham's School of Education.

BEST COPY AVAILABLE

FairTest Publications

______ ***FairTest Examiner***, published quarterly. . . $15.00/year (individuals), $25.00 (institutions).

______ Back Issues of ***FairTest Examiner***, $4.00 each: Vol ___No. ___(All 9 for $32.00)

______ ***Standing Up to The SAT***, by John Weiss, Barbara Beckwith and Bob Schaeffer (192 pp.) . . . $6.95

______ ***Fallout From the Testing Explosion: How 100 Million Standardized Exams Undermine Equity and Excellence in America's Public Schools***, by Monty Neill and Noe Medina (45 pp.). . . $8.95.

______ ***Sex Bias in College Admissions Tests: Why Women Lose Out***, by Phyllis Rosser and FairTest staff, 3rd ed. (60 pp.) $7.95

______ ***The Reign of ETS***, by Allan Nairn and Ralph Nader (550 pp.). . . $30

______ ***Beyond Standardized Tests: Admissions Alternatives That Work***, by Amy Allina with FairTest staff (18 pp.) $5.50

______ ***None of the Above, Behind the Myth of Scholastic Aptitude***, by David Owen (327 pp.) . . . $5.95

Updates, issued bi-annually (Spring & Fall). $12/year/topic

______**Testing Reform Initiatives** ______ **University Admissions Tests**

______**K-12th Grade Testing** ______ **Teacher/Employment Testing**

Contributions to FairTest

I agree there's no better way to be part of the testing reform movement than supporting FairTest! Enclosed is my tax-deductible contribution!*

______ **FairTest Associate** ($25.00--includes a free subscription to the *FairTest Examiner*)

______ **FairTest Sponsor** ($50.00--includes *Examiner* subscription and *Standing Up to the SAT*)

______ **FairTest Sustainer** ($100.00 or more--includes above plus a 20% discount on publications)

______ **Other**

Publications Total ______

Contributions Total ______

Check Total ______

Name ______

Address ______

City ______ State ______ 155 ZIP ______

Please make checks payable to FairTest and mail to: FairTest, 342 Broadway, Cambridge MA 02139

**Contributions to FairTest are fully tax deductible, except for the value of publications received.*

***Bunker Strategy Adopted in Response to* Standing Up to the SAT**

Book Sales Soar as Test-Makers Retreat

FairTest's new book, *Standing Up to the SAT* (ARCO/ Simon & Schuster), was published September 11, 1989 to a surge of critical acclaim and a rush of buying orders. Already in its second printing, the book has earned favorable media attention from coast-to-coast.

Meanwhile, the College Board and Educational Testing Service have retreated to their offices, alternately trying to ignore the issues raised in the book and unleashing false and misleading broadsides. For example, College Board spokesperson Fred Merino told the *Chronicle of Higher Education* that the test-takers' "Bill of Rights," a centerpiece of *Standing Up to the SAT*, is unnecessary since "students already *have* all the rights FairTest is calling for" (emphasis in original). But, as the book makes clear, students *don't* have rights such as "informed consent" before the administration of the SAT experimental section or "due process" when the test-maker claims that cheating has occurred.

If Mr. Merino were telling the truth and students already do have these rights, why did the College Board and ETS spend thousands of dollars in test-takers' fees over the past two years lobbying against New York State legislation containing just such protections? And why do cases like the one dramatized in the movie "Stand and Deliver" continue to occur (see page 11 for details.)

College Board President Donald Stewart has been even more intemperate, accusing critics like FairTest of "reverse racism" and "intellectual colonialism." Apparently, ETS has made no official statement about the book.

Since the book's publication, the College Board and the Educational Testing Service have backed away from debating the charges of test bias, inaccuracy and irrelevance raised in *Standing Up to the SAT*. Company public relations offices turned down joint appearances with FairTest on ABC-TV's "Good Morning America," Cable News Network's "Crossfire," and dozens of radio talk shows around the country. In addition, ETS and the College Board refused to participate in a panel discussion at the National Association of College Admissions Counselors (NACAC) convention, forcing organizers to cancel the scheduled event.

Fortunately, this "bunker strategy" has not prevented FairTest from continuing its public education campaign, including addressing more than one hundred people at an informal seminar at the NACAC conference.

If you don't yet have a copy, autographed editions of *Standing Up to the SAT* are available from FairTest for $6.95 postpaid, the same price as in bookstores. Place your order today using the form on page 15. Please call FairTest Publications Director Mike Friedman at (617) 864-4810 to discuss attractive discounts available for bulk purchases. (Members of the news media can still obtain review copies by calling ARCO at (212) 373-8946.)

FairTest
National Center for Fair & Open Testing
342 Broadway
Cambridge, MA 02139-1802
(617) 864-4810
(617) 497-2224 (FAX)

ADDRESS CORRECTION REQUESTED

Non-Profit
U.S. Postage
PAID
Cambridge, MA
Permit No. 53278

Moving? Return your label with new address

FairTest
EXAMINER
The National Center for Fair & Open Testing Vol. 4, No. 1 Winter, 1989-90 ISSN # 0898-2511

Coalition Calls for "Genuine Accountability"

Over three dozen national education and civil rights organizations and leaders have joined FairTest in calling on the Bush Administration and the nation's 50 governors to take the lead in "phasing out standardized, multiple-choice tests and replacing them with reliable, valid, and educationally appropriate alternatives."

CAMPAIGN FOR GENUINE ACCOUNTABILITY
CAMPAIGN FOR GENUINE ACCOUNTABILITY

Participants in the January 24 news conference announcing the Statement on Genuine Accountability. *From l. to r. Bob Witherspoon of the National Coalition of Title I/Chapter I Parents, Barbara Willer of the National Association for the Education of Young Children and FairTest's Monty Neill. Photo by Beverly Orr.*

In their *Statement on Genuine Accountability*, released at a January 24 National Press Club news conference in Washington, DC, the groups said, "A heavy emphasis on standardized multiple-choice tests as the primary measure for national educational goals will undermine many of the educational reforms which the governors and President Bush want to achieve... Standardized multiple-choice tests are not an adequate means to measure educational progress, nor are high test scores an appropriate educational goal. Educationally sound

(continued on page 6)

NCAA Eases "Prop. 42," Keeps Cut-Scores

Following a third change in eligibility rules for interscholastic sports in the past seven years, FairTest has offered its services to the National Collegiate Athletic Association (NCAA) to "assist in crafting [athletic] eligibility rules which are fairer and more accurate than the use of cut-off scores on the SAT and ACT."

At its January, 1990 convention in Dallas, the NCAA voted to allow need-based scholarships to be awarded to student athletes who don't attain a minimum SAT or ACT score. Last year the convention had barred even those forms of assistance by adopting its controversial Proposition 42, which would have, in effect, denied financial aid to student athletes who didn't meet a test score cut-off.

FairTest and other critics claimed that Prop. 42 would deny access to higher education to many African American and Hispanic students who were capable of doing college-level work, but did not perform well on biased and inaccurate college admissions tests. Even the test-makers—the Educational Testing Service and the American College Testing Program—objected to the use of cut-off scores to determine athletic eligibility.

(continued on page 11)

CONTENTS

FairTest Seeks Organizer......2
Boston Tests Challenged 3
Calif. Says "No" to Tests 3
Gender Bias Hearing 4
Funders Leaving NMSC........5
Truth-in-Testing Setback 5
Signers of "Statement on Genuine Accountability" 7
New NAEP Standards 7
"Authentic Evaluation" What is It? 8-9
Reports Revised 14
New York Scholarships 16

TM017177

FairTest®
National Center for
Fair & Open Testing
342 Broadway
Cambridge, MA 02139-1802
(617) 864-4810
(617) 497-2224 (FAX)

***FairTest Examiner*, Volume 4, #1**
Published February, 1990

Staff: Mike Friedman, Publications Director/Office Manager; Monty Neill, Associate Director; Bob Schaeffer, Public Education Director; Cinthia Schuman, Executive Director; Jamie Souder, Administrative Assistant; Sarah Stockwell, University Tests Coordinator.

Board of Directors: Denise Carty-Bennia, Esq., CUNY Law School, Chair; Dr. Pamela George, North Carolina Central University; Paul Pottinger, Goodsmith and Co.; William Robinson, Esq., Dean, D.C. School of Law; Bob Schaeffer, Treasurer; Steve Solomon, Esq., Proskauer, Rose, Goetz & Mendelsohn; Chuck Stone, *Philadelphia Daily News*; Ron Vera, Esq., Barbosa & Vera; John Weiss, Public Policy Initiatives, President.

National Policy Panel: Leonard Beckum, City University of New York; Dr. Beverly Cole, NAACP; Dr. Harold Dent, Association of Black Psychologists; Arnold Fege, National PTA; Dr. Howard Gardner, Harvard University; Dr. Asa Hilliard, Georgia State University; Rep. Matthew Martinez (D-CA); Deborah Meier, Central Park East Secondary School; Sara Melendez, University of Bridgeport; Aurelio Montemayor, Intercultural Development Research Association; Harry Pachon, National Association of Latino Elected and Appointed Officials; Alfredo de los Santos, Maricopa Community Colleges; Eleanor Smeal, Fund for a Feminist Majority; Theodore Shaw, Esq., NAACP Legal Defense Fund; Dr. Gail Thomas, Texas A & M University; Rep. Ted Weiss (D-NY).

Funding for FairTest's work comes from the Golden Rule Corporation, the Rockefeller Family Fund, the Ford Foundation, the Phillip M. Stern Family Fund, and the Aaron Diamond Foundation.

FairTest Seeks Field Organizer

FairTest is spearheading a national movement of parents, educators and citizen groups concerned about the impact of standardized testing on pre-kindergarten through grade 12 education. To facilitate the growth and development of this movement, we are looking for a Field Organizer who will work on a variety of tasks, including:

- **Technical Assistance**: The Field Organizer will provide information, training and strategic advice to parents, teachers and a broad range of education and civil rights organizations.
- **Advocacy**: The Field Organizer will play a major role in the conceptualization, development and implementation of testing reform campaigns.
- **Public Education**: The Field Organizer will develop materials on testing reform, speak at national, regional and local conferences and work with the media to encourage accurate and balanced coverage of issues related to standardized testing.

Candidates should have prior experience in organizing reform campaigns; a demonstrated commitment to improving public education; excellent writing, speaking, public policy and "people" skills; and the ability to work well independently. In addition, candidates must be willing to travel on a regular basis and enjoy working in diverse racial and ethnic communities. People of color are urged to apply.

Starting salary is in the $24,000 - $32,000 range, depending on experience. Benefits include three weeks vacation and health coverage.

Corrections

Due to an editing error, the cartoon on page 11 of the Fall, 1989 *FairTest Examiner* was not credited. It came from the *ASCD Update*, published by the Association for Supervision and Curriculum Development. We apologize for any inconvenience.

● ● ●

An *Examiner* reader in India, who provided information for an article in the Summer, 1989 *Examiner* has informed us that the statement "Upper Caste people, who make up only 24% of the Hindu population of the country, hold 89% of the higher level civil service jobs reserved for Hindus" is incorrect. Our reader informs us that there is no reservation of civil service jobs specifically for Hindus. We apologize for this misunderstanding.

Speaking Engagements

FairTest staff and board members can be heard speaking at the following conferences and meetings:

March 23: National Women & the Law Conference, Detroit

March 31: Southern Association on Children Under Six, Dallas

April 8: Education Writer's Association of America, Chicago

May 10: New England Center for Equity Assistance, Worcester, MA

K-12 Testing

Graduation Rule Challenged in Boston

The Boston School Committee has voted to postpone for one year a policy requiring students to score an 8.5 on the reading portion of the Metropolitan Achievement Test (MAT) to receive a diploma. If implemented this year, the policy would have denied graduation to 1066 seniors, 40% of the entire class.

Education reform groups, led by the Massachusetts Advocacy Center (MAC) sharply criticized the policy, charging that many students were tracked into courses that never allowed them to progress beyond eighth-grade level work. Thus, students who performed adequately in the classes to which they were assigned faced the denial of their diplomas. MAC also charged that teachers were being pressured to focus on test preparation rather than the curriculum.

In a letter to the school committee, FairTest noted that the MAT requirement violated the guidelines for test use established in the *Standards for Educational and Psychological Testing* published by the American Psychological Association. Those guidelines state that major decisions "should not automatically be made on the basis of a single test score," and that tests should not be used for such purposes as graduation requirements unless the student has had an opportunity to learn the content area covered by the test.

FairTest, MAC and other education groups urged the city to rely less on multiple-choice tests, and instead implement performance-based evaluation and related educational reforms.

Meanwhile, the Center for the Study of Testing, Evaluation and Educational Policy at Boston College is developing a set of assessments for the city that would deemphasize multiple-choice tests. Designed by Professors George Madaus and George Cawthorne, the new tests will require students to produce writing samples, read aloud, perform scientific experiments, and do other tasks connected to the curriculum. "This is the cutting edge," Madaus told the *Boston Globe* (12/2/89), "and if done correctly could really change the way we think about what kids can do." The new exams will be piloted in up to six schools as early as next September.

California Ed. Summit Says Stop Multiple-Choice Tests

"All multiple-choice tests should be eliminated," reads one key recommendation from California's recent "educational summit." Organized by the state's Department of Education, the December meeting was attended by 300 department leaders, administrators, teachers, parents, and community people. The summit was called to develop new strategies to further educational reform in California.

The recommendation on assessment also said, "The dominant methods and formats of testing not only do not support the kind of teaching and learning that the state and national curriculum reform movement calls for -- they are actually retarding that movement... An assessment system which measures student achievement on performance-based measures is essential for driving the needed reform to a thinking curriculum in which students are actively engaged."

The summit cited the California Assessment Program (CAP) direct writing assessment as a good example of authentic evaluation. "Direct assessment methods calling upon students to write, make oral presentations, and work with other students to solve real-world problems must also be implemented in math, history, and science as soon as possible."

"The primary purpose [of the new assessments]. . . should be to inform the teaching-learning process," the summit said.

Quote of the Month

Mr. Bush also said students "must do more than identify names on a multiple choice question. They must understand the generosity of Andrew Carnegie, the genius of Alexander Graham Bell and the heroism of Rosa Parks."

—President George Bush, as quoted in the *New York Times* September 29, 1989, speaking to the Education Summit with the nation's governors.

University Testing

Oregon Stresses GPA Over Tests

Contrary to the myth that large state university systems can't function without using the SAT as a screening device, the eight colleges and universities in the Oregon State System of Higher Education rely primarily on high school grade point averages (GPA) to select students. Students who have taken 14 core courses in high school and who meet a minimum GPA (determined by each college or university) are automatically admitted.

Even though no school in the system rejects more than 20% of its applicants, SAT or ACT scores are still required, but they are not considered for most students. Students who do fall below the minimum GPA requirement may be admitted on the basis of test scores, predicted college grade point average (a formula including test scores), or other factors. From the testing reform perspective, the next step would be to make SAT and ACT scores optional in the entire admissions process, which would more accurately reflect the role they play.

Testing Commission Holds Gender Bias Hearing

Last fall, the National Commission on Testing and Public Policy held a hearing on gender bias co-sponsored by the National Women's Law Center in Washington, DC. The session was one in a series of nationwide public forums held by the Commission on the effects of testing on such populations as Hispanics, blacks, Asians and Hawaiians, and Native Americans.

FairTest Executive Director Cinthia Schuman, Dr. Nancy Cole, Executive Vice President of the Educational Testing Service, and other test critics and researchers testified about the breadth, impact and causes of test bias against females, especially on college admissions exams. Schuman argued that colleges should not require the SAT for admission, due to the test's pervasive flaws, biases and its 65 point margin of error. Females consistently score lower than males on the SAT despite receiving higher grades in college, which is what the SAT is supposed to predict.

Cole admitted that some features of the SAT, such as its fast-paced format and rewards for guessing are good issues "to keep the heat on. If women can demonstrate skills and capabilities, but do so in somewhat different ways, we need to know that and be very sensitive to it in order that our tests....take account of it."

Dr. Linda Winfield, Principal Research Scientist at the Center for Research on Effective Schooling for Disadvantaged Students, told how she was not recruited by colleges because of her SAT scores, despite having grades in the top 10% of her class, taking Advance Placement courses and being in the honor society of her school. Even more discouraging, her guidance counselor said she would be "wasting her parents' money to go to college." Winfield ended up waiting five years before enrolling in college.

Other speakers and panelists included Phyllis Rosser, of the Equality in Testing Project, Isabelle Katz Pinzler of the American Civil Liberties Union Women's Rights Project, Carol Nagy Jacklin of the University of Southern California, Leslie Wolfe of the Center for Women Policy Studies, Sylvia Johnson of the School of Education at Howard University, and researcher Esther Diamond.

The Commission is expected to issue its findings and recommendations this Spring. Copies of the report on the hearing are available from Mary McLane, at the National Women's Law Center (202) 328-5183.

TWO WEEKS ON THE JOB
AND STILL
NO MULTIPLE CHOICE QUESTIONS!

University Testing

Corporate Funders Choose Alternative to National Merit

Responding, in part, to concerns about the use of standardized tests in the selection process for National Merit Scholarships, the Prudential Insurance Company of America and several other major corporations have shifted funds from the National Merit Scholarship Corporation (NMSC) to the Citizens' Scholarship Foundation of America, Inc. (CSFA), of St. Peter, Minnesota.

Prudential recently wrote to its employees explaining that it was replacing the National Merit Program with the CSFA scholarships at the same level of funding. The memo said, "[t]he screening process will, however, no longer eliminate applicants based solely on their performance on the PSAT...." Prudential officials also told FairTest they wanted broader scholarship selection criteria. The National Merit Scholarship Corporation makes awards only to students who score very high on the Preliminary Scholastic Aptitude Test/ National Merit Scholarship Qualifying Test (PSAT/NMSQT).

Corporations currently fund more than one half of all scholarships awarded by the National Merit Scholarship Corporation: in 1989 more than 55% of the $33.2 million "Merit" scholarships given out by NMSC came from corporations. Prudential, a major funder, provided 22 scholarships. According to CSFA, other corporate donors have transferred part of their scholarship monies or added CSFA scholarships.

CSFA said concern about the PSAT/NMSQT is an underlying reason for the switch for some firms. Many, like Prudential, want broader criteria for scholarship winners. Others value the fact that CSFA scholarships can be used at two-year colleges and vocational/technical schools. Scholarships administered by NMSC are limited to four-year institutions.

CSFA says it selects scholarship winners for companies based on a set of factors, including high school grade point average, class rank, work experience, community and school activities, test scores and recommendations. The few companies that must employ a screening device because there are many applicants generally use a combination of grade point average and class rank. Test scores are never used as the sole qualifying factor for any scholarship.

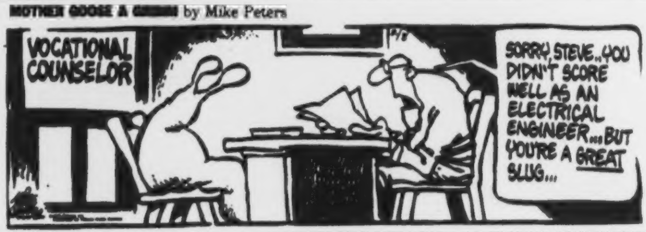
Prudential is the first corporation known to have dropped its funding from NMSC partially out of concerns about standardized testing. The coalition of FairTest, civil rights, women's rights and student groups opposed to the National Merit Scholarship Corporation's sole use of the PSAT/NMSQT to select semifinalists for National Merit Scholarships hopes other corporate funders will follow suit in the near future.

Truth-in-Testing Suffers Setback

In a blow to New York's ten-year-old Truth-in-Testing Act a federal district court has found that the landmark legislation violates federal copyright law. Truth-in-Testing protects students by requiring testing companies to publicly disclose university admissions test questions and answers, as well as make public research reports on test validity. While the decision applies only to the plaintiffs in the case—the American Association of Medical Colleges (AAMC), producers of the Medical College Admission Test (MCAT) —it leaves the door wide open for challenges or noncompliance by other testing companies.

In granting the AAMC's motion for summary judgment, Judge Neil P. McCurn, Chief Judge of the Northern District Court of New York, acknowledged that public disclosure of tests and test data serves strong public purposes, such as exposing tests' disparate impact on women and minorities. However, he decided that other factors were more important. The AAMC claimed, for example, that public disclosure of test questions

(continued on page 11)


MOTHER GOOSE & GRIMM by Mike Peters
VOCATIONAL COUNSELOR
SORRY, STEVE..YOU DIDN'T SCORE WELL AS AN ELECTRICAL ENGINEER...BUT YOU'RE A GREAT SLUG...

Reprinted with permission. Grimmy, Inc

K-12 Testing

Statement on Genuine Accountability

(from page 1)

means of evaluation exist and can provide a basis for improving teaching, informing the public, and measuring progress, thereby providing genuine accountability."

The Statement was initiated by FairTest working with the Council for Basic Education (CBE), the National Association for the Advancement of Colored People (NAACP), the National Association for the Education of Young Children (NAEYC), the National Parent Teacher Association, and the American Reading Council. Other signers include the American Federation of Teachers, the International Reading Association, and the National Association of Elementary School Principals (see opposite page for a complete list). The Statement was issued in an effort to dissuade the President and governors from relying on standardized tests to measure progress toward the educational goals espoused by the September, 1989 Education Summit.

In his statement to the press, FairTest Associate Director Monty Neill said, "The people who are on the front line of U.S. education—the groups that represent students, parents, teachers, curriculum specialists, principals, and those concerned with the education of low-income and minority-group children—all agree there is too much testing."

A. Graham Down, Executive Director of the CBE, explained, "The world we live in needs people who can think, write, evaluate, compare, and weigh competing ideas. These abilities are the essence of education. But they cannot be assessed by multiple-choice tests." He concluded, "Excellent techniques are available to replace multiple-choice exams."

The Statement explained that the alternatives to standardized tests "go by various names: authentic, appropriate, performance-based, and direct. They are evaluations based on asking students to demonstrate what we want them to be able to know and do" (see "What is Authentic Evaluation?," p.8-9).

Barbara Willer of the NAEYC said that her organization is "particularly concerned by the overemphasis on standardized achievement testing in many state reform efforts" because when tests are emphasized, "they inevitably drive the curriculum." She called for major changes in the education of young children, including changes in assessment.

Beverly Cole, Education Director of the NAACP, explained her organization's sponsorship in a written statement: "America needs measures of educational performance which provide a level playing field for all students. Standardized exams, with their well-documented biases and flaws, fail this test."

Speaking at the press conference, Bob Witherspoon, Executive Director of the National Coalition of Title I/ Chapter I parents, criticized the way in which standardized tests have come to dominate Chapter I programs. "Chapter I has become a test-driven program," he said, explaining that test scores are too often the sole basis for entry into the program. Once students are in the programs, he said, "They spend too much time on testing."

To further the joint effort exemplified in the Statement, FairTest is joining with a number of the signers to form the *Campaign for Genuine Accountability in Education*. Contact FairTest for more information.

FairTest Executive Director Cinthia Schuman speaking with A. Graham Down , Executive Director of the Council for Basic Education at the news conference announcing the Statement on Genuine Accountability. *Photo by Beverly Orr.*

K-12 Testing

Signers of "Statement on Genuine Accountability"

The following organizations and individuals have signed the "Statement on Genuine Accountability:"

American Association of Colleges for Teacher Education
American Federation of Teachers
American Reading Council
Angelo N. Ancheta, Asian Law Alliance*
APPLE Corps, Atlanta, GA
Association for Supervision and Curriculum Development
Association for Women in Science
Rims Barber, Mississippi Human Services Agenda
Leonard Beckum, City College of City University of New York
Diana Caballero, Puerto Rican/Latino Education Roundtable, NYC*
Center for Law and Social Policy
Center for Women Policy Studies
Council for Basic Education
Harold E. Dent, Psychological and Human Resources Consultants*
Education Law Center, Newark, NJ
Pablo Eisenberg, Center for Community Change*
Federation of Organizations for Professional Women
Fund for the Feminist Majority, Eleanor Smeal, President.
Howard Gardner, Project Zero, Harvard University*
Genesee Valley Developmental Learning Group, New York
La Donna Harris, Americans for Indian Opportunity*
Leslie A. Hart, Brain-Compatible Education Associates
Asa Hilliard, Georgia State University*
Holt Associates/*Growing Without Schooling*
Institute for Learning and Teaching, St. Paul, MN
International Reading Association
KEY-Kids, Education and You, Denver City, TX.
Massachusetts Advocacy Center
Matsushita Foundation, Inc.
Deborah Meier, Central Park East Secondary School, NYC*
Samuel J. Meisels, Univ. of Michigan*
Sara E. Melendez, Vice-Provost, University of Bridgeport*
Mexican American Women's National Association
National Association for the Advancement of Colored People
National Assocation for the Education of Young Children
National Association of Elementary School Principals
National Center for Fair & Open Testing (FairTest)
National Coalition of Title I/Chapter I Parents
National Council of Teachers of English
National Indian Youth Council
National Organization for Women—NYC Chapter
National Parent-Teacher Association
National Women's Law Center
National Women's Political Caucus
Fred M. Newmann, National Center on Effective Secondary Schools*
New York Public Interest Research Group (NYPIRG)
Vito Perrone, Harvard University*
Project Equality Inc., Kansas City, MO
Puerto Rican Legal Defense and Education Fund
Rochester (NY) Teachers Association
Lori Rubenstein, Partnership for Democracy, Washington, DC*
William V. Schipper
Donald H. Smith, Baruch College*
Southern Association on Children Under Six
Southern Christian Leadership Conference
Southern Regional Council
Gail E. Thomas, Texas A & M University*
United Federation of Teachers (NYC)
United States Student Association
Dorothy J. Watson, Whole Language Umbrella*
Paul Weckstein, Center for Law and Education*
Grant Wiggins, CLASS
Arthur E. Wise
Women's Research and Education Institute

*Organization listed for identification purposes only

New Standards on NAEP Tests?

In a statement to the National Assessment Governing Board (NAGB), FairTest urged the Board not to mandate country-wide age- and grade-level standards based on scores on the National Assessment of Educational Progress (NAEP). In invited testimony delivered at a Washington, D.C. hearing, FairTest Associate Director Monty Neill warned the board that if they responded to pressure for educational reform with a "quick fix: yet another test," the results would "most certainly cause harm to the nation's educational system."

In the wake of the "Education Summit" and a new NAEP report showing almost no gain in reading and writing from 1971 to 1988, NAGB held the hearing on whether to implement national performance standards on its 1990 math assessment. These tests will be administered to a nationwide sample of 4th, 8th and 12th-grade students. The proposed standards would use responses to selected questions as indicators of capability in various skill areas.

In other testimony at the NAGB hearing, Susan Harman of the American Reading Council explained that selected, narrow content reduced to items on a test could not be an adequate measure of student capabilities. This is true whether the narrow items are multiple-choice or open response. What is needed, she said, is documentation of real work by students. Graham Down of the Council for Basic Education also criticized

(continued on page 13)

K-12 Testing

What Is Authentic Evaluation?

Authentic evaluation of educational achievement directly measures actual performance in the subject area. Standardized multiple-choice tests, on the other hand, measure test-taking skills directly, and everything else either indirectly or not at all.

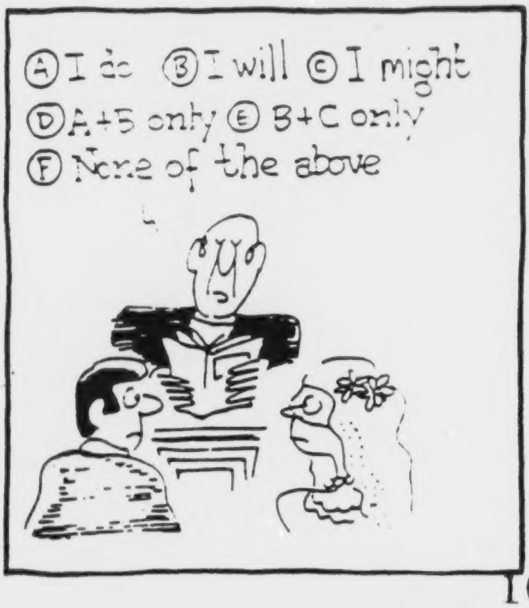
Also called "performance," "appropriate," "alternative," or "direct" assessments, authentic evaluation includes a wide variety of techniques, such as written products, solutions to problems, experiments, exhibitions, performances, portfolios of work and teacher observations, checklists and inventories, and cooperative group projects. They may be the evaluation of regular classroom activity or take the form of tests or special projects.

Authentic evaluations indicate what we value by directing instruction toward what we want the student to know and be able to do. They are appropriate to the student's age and level of learning and the subject being measured, and are useful to both teachers and students.

All forms of authentic assessment can be summarized numerically or put on a scale. Therefore, individual results can be combined to provide a variety of information about aggregate performance at the classroom, school, district, state and national levels. Thus, state and federal requirements for comparable quantitative data can be met.

Authentic evaluations indicate what we value by directing instruction toward what we want to student to know and be able to do.

Authentic assessment was developed in the arts and in apprenticeship systems, where assessment has always been based on performance. It is impossible to imagine evaluating a musician's ability without hearing her sing or play an instrument, or judging a woodworker's craft without seeing the table or cabinet. It is also impossible to help a student improve as a woodworker or musician unless the instructor observes the student in the process of working on something real, provides feedback, monitors the student's use of the feedback, and adjusts instruction and evaluation accordingly. Authentic assessment extends this principle of evaluating real work to all areas of the curriculum.



Reprinted with permission, Joel Pett

The most widely used form of authentic assessment in education today is in writing. For example, twenty-eight states, the National Assessment of Educational Progress (NAEP), and many other nations ask students to write on assigned topics. The essays and stories are graded by teams of readers (usually teachers) who assign grades according to standard guidelines. The readers are trained and retrained throughout the process to maintain reliable standards, a process that produces a high degree of agreement among judges. As with all the examples, this methodology can be used to evaluate classroom work that has been collected in a portfolio; it only has to be adjusted for subject area and student age.

Similar procedures are now being followed with open-ended mathematics questions. These ask students to write their own response to a problem. There is no single way to find a "right answer" because the question is

(Continued opposite)

K-12 Testing

Common Questions Answered

designed to see how a student thinks through a problem, thereby indicating her or his ability to use math. The answers are scored by groups of teacher-readers, again following a standard grading procedure. Two-sevenths of the NAEP math questions will be open-ended in 1990.

Performance assessments in science ask students to plan or perform experiments or use scientific apparatus, as is done in New York State. Science assessments can be graded by observation (where a teacher or other observer uses a checklist) or by scoring the students' written answers to the questions. These assessments can be designed to indicate understanding of basic scientific concepts and methods.

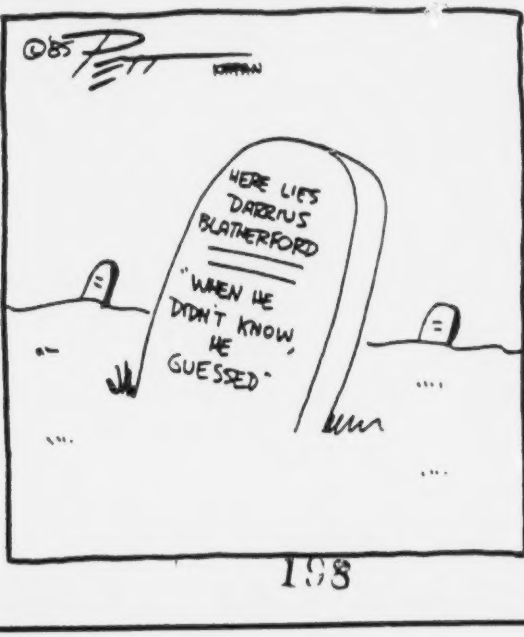
History/social studies assessments frequently require group projects, such as preparing a history of the neighborhood or discovering how a group of people changed a law or policy, tasks which allow students to demonstrate that they grasp important concepts about history and about democratic processes. Foreign language assessments ask students to use the language in a real-life situation, orally and in writing.

For young students, reading is best evaluated by having a student read aloud from material of varying levels of difficulty, while keeping a record of "miscues" that reveal the reader's strengths and weaknesses and the strategies used to solve problems. The reading can be taped and reviewed by teacher and student for further analysis and to monitor progress. The material can be discussed to evaluate comprehension and critical thinking. A writing assignment responding to the ideas of the reading passage can reveal the student's proficiency and thinking in both reading and writing.

All these assessments can be designed to closely follow the curriculum. They provide continuous, qualitative data to be used by teachers to help instruction. They can be used by students, who learn to assume responsibility for their portfolios and records and thereby engage in regular self-analysis of their work and progress. They provide a direct measure of achievement and, therefore, are worth the time spent preparing for and doing them. They also encourage an intelligent, rich curriculum rather than the dumbed-down, narrow curriculum fostered by teaching to and coaching for multiple-choice tests.

Teachers can and should be assisted in the evaluation process by community groups, parents, administrators, and university faculty. Outside participation can ensure that racial or cultural bias does not distort the assessment process. For example, a team can examine student portfolios and then compare their evaluations with those of the teacher. These teams should also be helpful in strengthening the evaluation capabilities of teachers by providing feedback.

Authentic evaluation will provide far more information than any multiple-choice test possibly could. The costs of teacher involvement in designing, administering, and scoring new assessments can be counted as part of professional and curriculum development, since no other activity involves teachers more deeply in thinking about their teaching, its objectives, methods and results. Schools and communities will see that authentic assessments are promoting the thinking curriculum everyone wants for our children, and thereby providing genuine accountability.


©85 PETT
HERE LIES DARRIUS BLATHERFORD
"WHEN HE DIDN'T KNOW, HE GUESSED"

Reprinted with permission, Joel Pett

Contemporary Living

The Atlanta Journal | THE ATLANTA CONSTITUTION ••••• Tuesday, October 24, 1989

Book Urges Students To Stand Up to SAT

By Rebecca McCarthy
Staff Writer

Remember those nail-biting Saturdays spent with two stiletto-sharp No. 2 pencils and the Scholastic Aptitude Test (SAT)? The standardized exam looms on the horizon for thousands of high school juniors and seniors, many of whom have registered to take it during the next three months.

They'll find help in bookstores: a slim paperback called "Standing Up to the SAT" (ARCO, $6.95). The book claims to give readers the mental muscles for overcoming what it calls "the basic flaws and biases" of the SAT, which is used in the admissions process of more than 1,600 U.S. colleges.

"We're telling students, 'You can beat the test, you don't have to let it beat you,' " says Bob Shaeffer of FairTest, a non-profit consumer advocacy group in Cambridge, Mass. He co-wrote the 186-page book. "We want them to realize they have a right to be judged on a real performance, not on three hours they spend taking a test."

More specifically, the book brands the SAT "a culturally narrow test that systematically discriminates against lower-income, minority, female, bilingual and

ADVICE Continued on **G2**

Advice on 'Standing Up to the SAT'

From Page G1

rural students."

Created in 1926, the SAT measures verbal and math skills through multiple-choice questions. It is sponsored by the College Board, a non-profit consortium of high schools and colleges, and is administered by the Educational Testing Service, a Lawrenceville, N.J.-based company.

Though they have declined to publicly debate the book's criticisms, College Board officials did fire off a four-page press release citing the authors' "unsubstantiated allegations" and dismissing the book as a "misleading ... rehash of old charges that the Board, researchers, educators and other professionals have answered before."

Mr. Shaeffer agrees, to a point. "Many of these criticisms have been raised before — but the College Board has never answered them."

While debate continues on the merits of both book and test, sales of "Standing Up to the SAT" have been brisk since it was published last month; 25,000 copies have already been sold. "It contains advice that students needed and deserved," says ARCO publisher Pam Byers. "We're not saying that standardized testing is terrible, but that students should have the benefit of this practical information."

ARCO
IN ANY OTHER

'Standing Up to the SAT' claims to help readers cope with 'basic flaws and biases' in the exam.

In its first few chapters, the book presents a 10-point SAT Bill of Rights designed to de-mythologize the exam and to bolster self-confidence.

It includes the right to a test that measures what it claims to measure; to be fully informed about how to prepare; to privacy for scores and other personal data; to due process for any challenge to the test; and to open and accountable testing procedures. College Board officials contend such rights have existed for years.

Targeted at teenagers — and their parents — the book offers step-by-step information about registering for the SAT, selecting a coaching course (for those who can afford it) and setting up a test-preparation schedule well in advance of the exam date. Students are urged to study previous SATs, to skip time-consuming questions and to make intelligent guesses because "the 'wanted' answer is right in front of you, surrounded by several 'distractor' choices."

"Standing Up to the SAT" concludes by "telling students how to work to change the public perception about standardized testing," Mr. Shaeffer says. "We want them to learn from their experience and join with FairTest to help protect other students from having to take similarly biased tests."

BEST COPY AVAILABLE

University Testing

Truth-in-Testing

(from page 5)

would prevent their reuse, thus exhausting a limited pool of test items and harming the marketability of the exam.

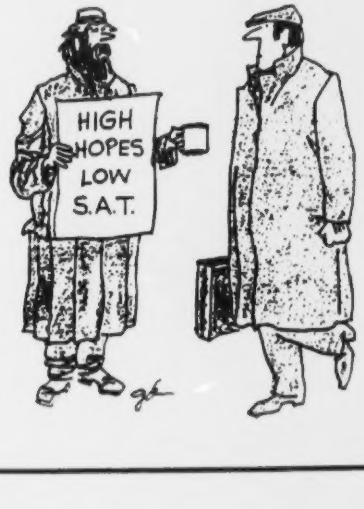
Ironically, arguments such as these were made by test manufacturers in the late 1970s when they lobbied against the bill. They also claimed that Truth-in-Testing would result in much higher costs for test-takers, the disruption of test development and poorer quality tests. None of these dire predictions came true.

In fact, soon after the law went into effect in 1980, the manufacturers of the Scholastic Aptitude Test (SAT), American College Testing Program Assessment (ACT), Graduate Record Examination (GRE), Law School Admissions Test (LSAT) and Graduate Management Admissions Test (GMAT) voluntarily extended the law's provisions nationally. They did this in the face of Truth-in-Testing bills pending in over 20 state legislatures and the U.S. Congress.

Since then, implementation and acceptance of the concept has gone smoothly. As a result of Truth-in-Testing, hundreds of students have been able to challenge test questions and thousands of scores have been raised. In addition, the measure has enabled watchdog groups such as FairTest to monitor testing companies and critique their products. Even the Educational Testing Service (ETS)—which spent over $30,000 lobbying against the bill—has admitted that Truth-in-Testing is, in general, "a positive concept."

Still, ETS and ACT have not ruled out the possibility of using the court decision to restrict Truth-in-Testing rights. ETS general counsel Stanford Mayrhauser told the *New York Times*, "If you're looking for a categorical denial that we won't seek to avail ourselves of any remedies made available to us by this ruling, I won't deny that" (*New York Times*, January 19, 1990).

The AAMC, on the other hand, has never complied with the Truth-in-Testing law. In 1980, they were successful in obtaining a preliminary injunction against the enforcement of Truth-in-Testing for the MCAT. The case had been pending for over a decade. The State of New York is reviewing the decision to determine whether it will appeal.



Reprinted with permission

Prop. 42

(from page 1)

A FairTest memorandum, mailed to more than one thousand delegates to the 1990 convention, called upon the NCAA to "support fair and relevant athletic scholarship criteria by voting to overhaul Proposition 42." The memo challenged the NCAA to produce any data demonstrating that Prop. 42 was academically justified. It concluded that Prop. 42 was a blatantly improper use of test scores.

Criticism of Proposition 42 has been widespread. The Hawaii legislature adopted a resolution requesting the President of the University of Hawaii to urge the NCAA to repeal Prop 42. The New York Assembly passed a resolution calling on ETS to withhold scores from the NCAA. In addition, the Black Coaches Association issued a scathing statement, condemning Prop. 42, before the convention. The American Civil Liberties Union and the New York Public Interest Research Group sent letters to NCAA delegates protesting the misuse of test scores in NCAA regulations.

The NCAA's decision to eliminate the worst features of Prop. 42 avoided the fundamental issue of whether college admissions tests should be used at all to determine athletic eligibility. Under the 1990 rule (Prop. 26), student athletes who fall below the test score cut-offs will still be ineligible for athletic scholarships and can't play in their first year. But financial aid—and, for many, the opportunity for a college education—will no longer be denied simply because of test scores.

The Boston Globe, Sunday, January 28, 1990

Reprinted with permission

MURIEL COHEN

Standardized tests challenged again

LAST WEDNESDAY, A broad sweep of influential political, civil rights and educational organizations appealed to President Bush, the Congress, state legislatures and the 50 governors to drop standardized tests as a measure of the effectiveness of schools and the performance of students.

A statement issued by FairTest, the Cambridge-based watchdog group, and signed by representatives of more than 30 agencies contended that teacher observation, student portfolios and exhibits are more accurate and reliable assessments than the tests now in use.

It was yet another example of David taking on Goliath.

Once again, FairTest has taken on the big test companies in an arena where the stakes run in millions of dollars.

The industry giant is Educational Testing Service, which prepares a number of tests and administers many of them, including the Scholastic Aptitude Test, to millions of students annually. ETS president Gregory R. Anrig, responding to the FairTest statement, said last week, "I cannot agree with your blanket rejection of multiple-choice tests. They remain a valid, efficient and inexpensive means to measure certain important aspects of student achievement."

At the same time, Anrig said that ETS is developing new measures to help schools evaluate students. Anrig said that Harvard professor Howard Gardner, one of the signers of the FairTest appeal, is working with ETS in Pittsburgh on developing portfolio assessment of student work in art, music and creative writing.

When he was Massachusetts commissioner of education, Anrig refused to approve the use of standardized test results to enable comparison of schools or school districts. Such testing was initiated in Massachusetts under his successor, John Lawson.

In its latest attack on tests, FairTest has assembled some powerful allies, including the American Federation of Teachers, the National Association for the Advancement of Colored People, the National Association for the Education of Young Children and many more.

FairTest is hammering away at a problem widely acknowledged by educators and policy-makers – that such tests, such as the Metropolitan Achievement and the California Basic Skills – skew curriculum, aggravate student anxieties and are frequently used for accountability, rather than as a guide to improving instruction.

"It is crucial that the governors, the administration, Congress, state legislatures and other policy and education officials recognize that standardized, multiple-choice tests are not an adequate means to measure educational progress, nor are high test scores an appropriate educational goal," said FairTest.

FairTest describes alternatives, such as "asking students to demonstrate what we want them to be able to know and do. These assessments provide instructionally useful information, can avoid the cultural and racial biases of many standardized tests and can be put into a form that will provide useful data and genuine accountability to local, state and national governments."

FairTest's protest comes four weeks before the National Governors Association meeting, which is expected to produce national education goals. The NGA has been working with White House aides to formulate an education platform as recommended at President Bush's education summit last September.

The "Statement on Genuine Accountability" prepared by FairTest said, "A heavy emphasis on standardized multiple-choice tests as the primary measury for national educational goals will undermine many of the educational reforms which the governors and President Bush want to achieve."

In this instance, FairTest is right on target. There is a basic contradiction in all of the national reform cries urging school-based reforms while insisting on standardized testing.

Tests should be used for diagnosing student weaknesses and as guides for remedies, not to produce data for political use – or abuse.

NAEP

(from page 6)

narrow and reductionist approaches that limit schooling to multiple-choice test preparation.

Keith Geiger, President of the National Education Association, recomended that NAGB wait until the new assessment had proven itself before setting standards. Michael Cohen of the National Governors' Association (NGA) also urged NAGB to wait because the Governors, with whom NAGB wants to work, had not yet decided on the goals they would support.

NAGB, which oversees the NAEP exams, plans to decide the issue at its March meeting and then work with the Bush Administration and the NGA to determine what capabilities should be measured.

Few Gains on NAEP Tests

The 1980s were a decade of massive explosion in the quantity of standardized tests and in their effects on curriculum. However, recent NAEP reports on reading and writing indicate there has been little improvement in students' abilities to read or write. Students at all levels showed no gains in their ability to perform higher order tasks, such as understanding and synthesizing complex material.

The reading report also showed that the measured capabilities of nine-year-olds declined. FairTest told NAGB, "Children whose early years of schooling occurred in the test-dominated 1980s are not doing as well as children whose primary school years were the 1970s. School reform schemes directed primarily toward improving performance on multiple-choice-type tests have not and will not lead to meaningful improvements in student achievement."

K-12 News in Brief

New Georgia Assessment

The state of **Georgia** is developing an assessment program that state officials hope will foster developmentally appropriate schooling for kindergartners, that is, schooling which understands the processes and variety of children's social, emotional and intellectual growth. "The glory of this instrument is that it is an instructional tool, not just for children but for the adults who administer it," Diane Cousineau explained to *Education Week* (11/15/89). Last spring, the state dropped its controversial, mandated use of the California Achievement Test as an entry hurdle for first grade.

Gesell in Trouble

According to a report in *Education Week* (10/25/89), the **Gesell Institute**, maker of the Gesell School Readiness Tests, is facing financial difficulty and an "uncertain future." The Gesell tests have been frequently criticized as lacking reliability and validity evidence; having been normed on a tiny, mostly-white, Connecticut sample; being based on outdated psychological theories; and encouraging inappropriate educational practices.

Jersey Report Cards Blasted

New Jersey's recently issued statewide school "report cards" have been criticized for promoting flawed comparisons among schools based on simplistic use of test scores. These comparisons have ignored income and per-pupil spending differences among communities when reporting scores. The state's own analyses have shown that the higher a district's income, the higher the test scores, but the state omitted socio-economic data from the "report cards" (*N.Y. Times*, 11/27/89). Controversy over the inclusion of SAT averages led the New Jersey Principals and Supervisors Association and Assemblyman Anthony Impreveduto to file legislation prohibiting the use of admissions tests scores on the report card. Send an SASE to FairTest for a copy of the bill.

ETS National Assessment

The **Educational Testing Service** (ETS) is developing a national system to assess the skills of high school graduates and make the results available to employers. The plan, named "Worklink," has initial support from the American Business Conference and the National Alliance of Business. Part of the plan will be a test, but an ETS spokesman said it would "use teacher ratings and other kinds of assessments." In response, Gordon Cawelti, of the Association for Supervision and Curriculum Development, said, "The last thing in the world schools need is one more test." Ted Sizer of the Coalition of Essential Schools suggested that business ought to help reform education, not promote more tests. *(Education Wk.* 11/15/89.)

New Book Exposes Testing

The **National Association for the Education of Young Children** has just released *Achievement Testing in the Early Grades: The Games Grown-Ups Play*, edited by Constance Kamii. This is a very readable introduction to the problems caused by standardized achievement tests and solutions to the problems. It contains chapters directed toward teachers, superintendents, and other levels of the education system, and provides information on appropriate assessment in literacy and math. Available from NAEYC, 1834 Connecticut Ave., N.W., Washington, D.C. 20009; $8.00.

FairTest Publications

Revised Edition of *Fallout* On The Way

In March, FairTest will release a revised edition of its groundbreaking report, *Fallout from the Testing Explosion*. This document caught the nation's attention by reporting that over 100 million standardized tests are administered yearly to U.S. schoolchildren. *Fallout* helped to focus the national debate on the overuse and misuse of multiple-choice type tests and point to assessment alternatives that are being implemented in many states and school districts.

The revised edition of *Fallout* will provide further analysis on how these exams harm students, particularly those from low-income and minority-groups, through tracking, retention in grade, and separating students into different programs. It will also provide further evidence on how tests narrow and distort curriculum and instruction, as well as more information on the theories behind test construction. In addition, the annotated bibliography has been significantly revised and enlarged.

You can order the forthcoming edition of *Fallout from the Testing Explosion* at the same low price of $8.95 by filling out the form on the opposite page.

Fact Sheets Available

FairTest has published three new fact sheets on testing from pre-school through twelfth grade. The first, *What's Wrong With Standardized Tests*, answers common questions about the construction of tests, their inability to adequately measure ability or achievement, and biases in the exams.

The second, *How Standardized Testing Damages Education*, responds to questions about the ways test misuse harms students and education. The contents of the third fact sheet, *What is Authentic Evaluation?* are reprinted on pages 8 and 9 of this issue.

Each fact sheet is available on one two-sided sheet of paper in a format useful for copying and distributing widely. For your free fact sheets, send a long $.25 SASE to: K-12 Fact Sheets, FairTest, 342 Broadway, Cambridge, MA 02139.

Revised SAT Bibliography

FairTest has recently updated and expanded its annotated bibliography on SAT bias and misuse. For a copy of the 14 page bibliography, which lists nearly five dozen sources, send your request along with 85 cents in stamps to FairTest, 342 Broadway, Cambridge, MA 02139-1802.

New Edition of *Sex Bias* Report Issued

FairTest's *Sex-Bias in College Admissions Tests: Why Women Lose Out*, the report that made sex-bias in university admissions exams a national issue and launched the campaign to reform the National Merit Scholarship awards process, has been reissued in a revised edition. The report, by Phyllis Rosser and the FairTest staff, now includes a foreword reviewing major public policy and research developments since the spring of 1987, when the report was originally published.

Appendices include "FairTest's SAT Bill of Rights," the most recent score averages for the SAT and ACT, a 1988-89 state-by-state chart of National Merit Semifinalists by gender, and a public letter to the National Merit Scholarship Corporation from FairTest and nearly 50 civil rights, feminist, student and education organizations. A selection of newsclips highlights the New York scholarship lawsuit, NCAA Proposition 42, and the books *The Case Against the SAT* and *The SAT Gender GAP*.

Even with all this new material, the report still costs only $7.95. To order, please use the order form on page 15.

Order FairTest Publications Today

☞

FairTest Publications

______ ***FairTest Examiner***, published quarterly. . . $15.00/year (individuals), $25.00 (institutions)

______ Back Issues of ***FairTest Examiner***, $4.00 each: Vol.   No.   (All 11 for $32.00)

______ ***Standing Up to The SAT***, by John Weiss, Barbara Beckwith and Bob Schaeffer (192 pp.) . . . $6.95

______ ***Fallout From the Testing Explosion: How 100 Million Standardized Exams Undermine Equity and Excellence in America's Public Schools***, by Noe Medina and Monty Neill (45 pp.). . . $8.95

______ ***Sex Bias in College Admissions Tests: Why Women Lose Out***, by Phyllis Rosser and FairTest staff, 3rd ed. (60 pp.) $7.95

______ ***The Reign of ETS***, by Allan Nairn and Ralph Nader (550 pp.). . . $30

______ ***Beyond Standardized Tests: Admissions Alternatives That Work***, by Amy Allina with FairTest staff (18 pp.) $5.50

______ ***None of the Above, Behind the Myth of Scholastic Aptitude***, by David Owen (327 pp.) . . . $5.95

Contributions to FairTest

I agree there's no better way to be part of the testing reform movement than supporting FairTest! Enclosed is my tax-deductible contribution!*

______ **FairTest Associate** ($25.00--includes a free one-year subscription to the *FairTest Examiner*)

______ **FairTest Sponsor** ($50.00--includes *Examiner* subscription and *Standing Up to the SAT*)

______ **FairTest Sustainer** ($100.00 or more--includes above plus a 20% discount on publications)

______ **Other**

Publications Total ______

Contributions Total ______

Check Total ______

Name ______

Address ______

City ______ State ______ ZIP ______

Please make checks payable to FairTest and mail to: FairTest, 342 Broadway, Cambridge MA 02139

**Contributions to FairTest are fully tax deductible, except for the value of publications received.*

Call For End to All SAT Use in NY Scholarships

More than one year ago, Federal District Judge John Walker barred New York State from using SAT scores alone to award up to $40 million annually in Regents and Empire Scholarships. In its landmark ruling, which found that the State's use of the SAT was unconstitutional and violated federal equal education laws, the court concluded that relying on the SAT discriminated against women because test scores "do not mirror high school Math and English classes" and fail "to provide *any* measure of what a student learns in foreign language, science, and social studies courses." Walker also found that the SAT underpredicts for females, saying "the probability, *absent discriminatory causes*, that women would consistently score 60 points less on the SAT than men is nearly zero" (emphasis added).

Reacting to the court order, the NY State Education Department (SED), which had defended the "SAT-only" method of selecting scholarships in court, designed a replacement formula using an extremely narrow interpretation of the court ruling. The SED's formula combined SAT scores with high school grade point average.

Though this new system increased the percentage of female winners, girls and minorities are still shortchanged. That is because the SED formula, in effect, weighs SAT scores, with all their biases and flaws, at least twice as heavily as grades.

A coalition of groups has formed to press the state to design a truly fair scholarship formula. Led by the New York Public Interest Research Group (NYPIRG), the New York Civil Liberties Union (NYCLU), and FairTest, the coalition has called for scholarships to be awarded according to classroom performance on a school-by-school basis. Using grades or class rank is a better measure of high school achievement, the coalition says, and would assure that women and members of minority groups receive a fairer share of scholarships.

A new report by NYPIRG, with the staff of FairTest, *Stopping the Scholarship Scam: Ending Gender and Race Discrimination in New York State's Scholarship Awards*, demonstrates that the current scholarship formula, in effect, earmarks a preponderance of awards for white males because it uses scores from an exam that discriminates against women and minorities. The NYPIRG analysis concludes that using three years of high school performance instead of scores from a three hour, multiple-choice exam will help, rather than hinder, educational reform.

Stopping the Scholarship Scam is available for $5.00 from NYPIRG, 9 Murray Street, New York, NY, 10007.

FairTest®
National Center for Fair & Open Testing
342 Broadway
Cambridge, MA 02139-1802
(617) 864-4810
(617) 497-2224 (FAX)

ADDRESS CORRECTION REQUESTED

52

Non-Profit
U.S. Postage
PAID
Cambridge, MA
Permit No. 53278

Paula Hannaman
American Institutes for Research
ERIC/TME
3333 K St. NW Suite 200
Washington, DC 20007-

3

Moving? Return your label with new address

FairTest EXAMINER

The National Center for Fair & Open Testing Vol. 4, No. 2 Spring, 1990 ISSN # 0898-2511

Testmakers File Suit to Overturn Truth-in-Testing

Advocates Assail ETS/CEEB Lawsuit

Consumer advocates, including FairTest, the New York Public Interest Research Group (NYPIRG) and Ralph Nader, have strongly criticized the Educational Testing Service (ETS) and the College Board (CEEB) for seeking to exempt their admissions exams from New York's landmark Truth-in-Testing law.

On April 16, the test-makers requested a preliminary injunction in federal court to administer exams in New York without the questions and answers becoming available to test-takers, as the 1979 law requires. In addition, the suit attempts to strike down the requirement that test-makers provide the state with research reports on test validity and fairness. In their press release announcing the suit, the College Board claimed that the injunction would "serve the public interest," since it would result in additional exams being given in New York.

"Once again, the nation's largest test-makers have demonstrated their contempt for public accountability and independent scrutiny of their products," said FairTest Executive Director Cinthia Schuman. "Secret exams increase the possibility of biased questions and scoring errors. Under the guise of providing more test-taking opportunities, the companies are attempting to obliterate protections students now have."

If successful in their lawsuit, the test-makers maintain that they would add new test dates when exams would not be disclosed, while voluntarily disclosing most university admissions tests in New York. For example, New York would have two additional administrations of the Scholastic Aptitude Test (SAT) and the Graduate Record Exam (GRE), and five additional administrations of the Test of English as a Foreign Language (TOEFL), all of which would be secret.

According to NYPIRG legislative director Blair Horner, "After New

(Continued on page 11)

NAEP Expansion Criticized

Educators around the country have stepped up criticism of the National Assessment Governing Board (NAGB), the body governing the National Assessment of Educational Progress (NAEP), charging it with undermining democratic processes and seeking to impose unwarranted federal control over education. NAGB's members were appointed by former Secretary of Education, William Bennett. NAEP periodically assesses a national sample of students in grades 4, 8 and 12 in a variety of subjects.

Recently, NAGB proposed expanding NAEP to district and school-level testing, which would allow test score comparisons down to the school building and possibly classroom levels. States would be allowed to require local school district participation; under current law, local participation is voluntary. Under the expansion proposal, the federal government would pay for all NAEP testing, an amount the U.S. Department of Education estimated at $100 million. The 1990 math and 1992 reading assessments are to include trial runs for developing

(Continued on page 6)

CONTENTS

New FT Board Member 2
Tracking in Boston 3
New NAEP Exams 4
Control of National Assessment Debated 5
SAT Optional 9
Truth-in-Testing Appeal ..11
Publications Order Form 15
Bar Exam Criticized 16

TM017189

FairTest®
National Center for
Fair & Open Testing
342 Broadway
Cambridge, MA 02139-1802
(617) 864-4810
(617) 497-2224 (FAX)

***FairTest Examiner*, Volume 4, #2**
Published April, 1990

Staff: Mike Friedman, Publications Director/Office Manager; Monty Neill, Associate Director; Bob Schaeffer, Public Education Director; Cinthia Schuman, Executive Director; Jamie Souder, Administrative Assistant; Sarah Stockwell, University Tests Coordinator.

Board of Directors: Denise Carty-Bennia, Esq., CUNY Law School, Chair; Dr. Pamela George, North Carolina Central University; Dr. Susanna Navarro, California Achievement Council; Paul Pottinger, Goodsmith and Co.; William Robinson, Esq., Dean, D.C. School of Law; Bob Schaeffer, Treasurer; Steve Solomon, Esq., Proskauer, Rose, Goetz & Mendelsohn; Chuck Stone, *Philadelphia Daily News;* Ron Vera, Esq., Barbosa & Vera; John Weiss, Public Policy Initiatives, President.

National Policy Panel: Dr. Leonard Beckum, City University of New York; Dr. Beverly Cole, NAACP; Dr. Harold Dent, Association of Black Psychologists; Arnold Fege, National PTA; Dr. Howard Gardner, Harvard University; Dr. Asa Hilliard, Georgia State University; U.S. Rep. Matthew Martinez (D-CA); Deborah Meier, Central Park East Secondary School; Sara Melendez, University of Bridgeport; Aurelio Montemayor, Intercultural Development Research Association; Harry Pachon, National Association of Latino Elected and Appointed Officials; Alfredo de los Santos, Maricopa Community Colleges; Eleanor Smeal, Fund for a Feminist Majority; Theodore Shaw, Esq., NAACP Legal Defense Fund; Dr. Gail Thomas, Texas A & M University; U.S. Rep. Ted Weiss (D-NY).

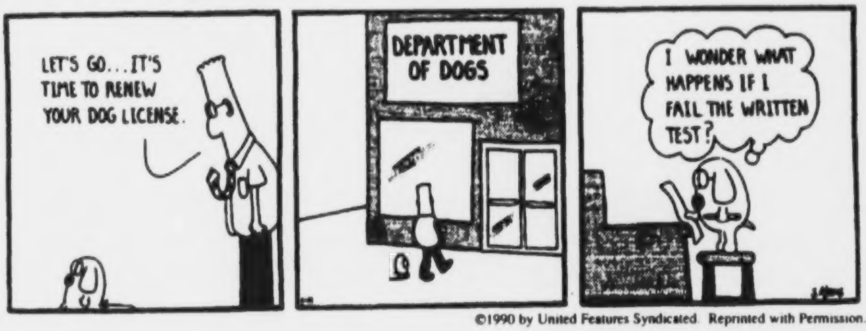
Funding for FairTest's work comes from the Golden Rule Corporation, the Rockefeller Family Fund, the Ford Foundation, the Phillip M. Stern Family Fund, and the Aaron Diamond Foundation.

FairTest Elects New Board Member

At its April 1990 meeting, the FairTest Board of Directors elected Dr. M. Susana Navarro, as its tenth member. Dr. Navarro has been Executive Director of the California Achievement Council since September 1989, and before that was the Council's Associate Director. The Achievement Council is a statewide advocacy group established by business leaders and the University of California to encourage families to become more involved in their children's educational achievement.

Dr. Navarro holds a Ph.D in Educational Psychology from Stanford University. While there, she received a Ford Foundation Fellowship. Dr. Navarro also worked at the Mexican American Legal Defense and Educational Fund (MALDEF), in San Francisco, where she worked with current FairTest board member Ron Vera on testing and other education issues. In addition to consulting for both the Educational Testing Service and the College Board, Dr. Navarro served on the ETS Visiting Committee in 1987. FairTest's welcomes Dr. Navarro and hopes that her association with the organization will be long and productive.

Dilbert, by Scott Adams


LET'S GO... IT'S TIME TO RENEW YOUR DOG LICENSE.
DEPARTMENT OF DOGS
I WONDER WHAT HAPPENS IF I FAIL THE WRITTEN TEST?

©1990 by United Features Syndicated. Reprinted with Permission.

Quote of the Month

"All multiple-choice tests should be eliminated. The current approach to assessment of student achievement which relies on multiple-choice student response to a narrow scope of desired student outcomes must be abandoned because of its deleterious effect on the educational process. An assessment system which measures student achievement on performance-based measures is essential for driving the needed reform to a thinking curriculum in which students are actively engaged and are successful in meeting expectations for transition to various goals beyond high school."

--Meeting the Challenge, *the final report of the California Education Summit (December 1989)*

K-12 Testing

Tracking and Testing in Boston Blasted

Locked In/Locked Out: Tracking and Placement Practices in Boston Public Schools critically analyzes the educationally destructive effects of tracking in Boston's schools. According to the new report, prepared by the Massachusetts Advocacy Center (MAC), tracking and placement decisions in Boston are frequently made on the basis of a single standardized test score. Thus, says the report, testing plays an important role in undermining educational equity and quality through systemic tracking.

For example, students are placed in grade four "advanced work classes" based on their grade two standardized reading test scores. Children who are not in advanced work classes for several years are unlikely to be adequately prepared to score high on the Secondary School Admissions Test (SSAT), a requirement for entry into the city's high quality examination schools. In turn, students who do not attend examination schools are far less likely to enroll in college, particularly four-year colleges.

Thus, performance on a one-time, multiple-choice reading test in second grade plays a major role in determining a student's life chances. Moreover, readiness screening in Boston, occurring before grade two, can lead to in-class groupings and placements in transitional programs that largely define who does well on the grade two reading test.

The MAC report argues that the tests Boston uses for tracking are both educationally narrow and culturally biased. Boston compounds its tracking problem by regular reliance on standardized tests for promotion and retention decisions. Students who have been retained-in-grade (flunked) are three times as likely to drop out of high school as non-retained students. At all grades, those who are retained are also apt to be put in lower tracks.

The MAC report details the very different qualities of education offered students in different tracks, from advanced work to regular, bilingual and special education placements. For all but those at the top, schooling in Boston is often reduced to remediation, which tends to rely on repetitious drill aimed at raising test scores. As a result, only a small percentage of students have access to the educational knowledge valued in our society.

This system creates different outcomes for students from different racial and economic backgrounds. Whites and some Asians are, in general, in the top tracks. African Americans, Hispanics and other Asians—the great majority of Boston's public school students—are assigned to the bottom. The physical separation of top tracks from lower groups means that many schools become functionally segregated, all in the name of "ability grouping."

The MAC report calls for the abolition of this tracking-based system. It presents voluminous research information on the dangers of tracking and retention as well as many examples of effective schooling that does not utilize extensive tracking. The report concludes with a plan for how Boston can change tracking and with it the kinds of education offered to all children.

MAC is well aware of the obstacles to reform. First on its list is standardized testing, which perpetuates existing inequities and reinforces the wrong educational goals. School improvement, concludes the report, cannot be identified with increasing test scores.

This well-documented, 200-page report is available for $10 pre-paid from Massachusetts Advocacy Center, 95 Berkeley St., Boston, MA 02116.

Robotman, by Jim Meddick

I'M STUDYING TO RENEW MY FLYING SAUCER LICENSE. DO ME A FAVOR AND QUIZ ME...
"YOU'RE IN A VEHICLE WHICH HAS A MAXIMUM SPEED OF WARP SIX. A GIANT METEOR IS GAINING ON YOU AT WARP NINE. WHAT DO YOU DO?"
YOU SAY, "CAP'N KIRK, ME ENGINES CAN'T TAKE MUCH MORE O' THIS!" THEN YOU ACCELERATE TO WARP 10.
CORRECT.

©1990 by United Features Syndicated. Reprinted with Permission.

K-12 Testing

New NAEP Exams in Reading, Writing & Math

The National Assessment of Educational Progress (NAEP) recently conducted its 1990 Math assessment, and the National Assessment Governing Board (NAGB) has approved the design features for its 1992 NAEP Reading and Writing exams. The math and reading tests include NAEP's first state-by-state comparisons. The new assessments, however, have already come under fire.

The reading test will include three categories of reading. The time allotted for items has been expanded to allow for more open-ended questions as well as longer reading passages taken from real materials rather than materials produced solely for the test. Special studies on Oral Reading and Response and on Reading Portfolios, both at grade four, will also be conducted.

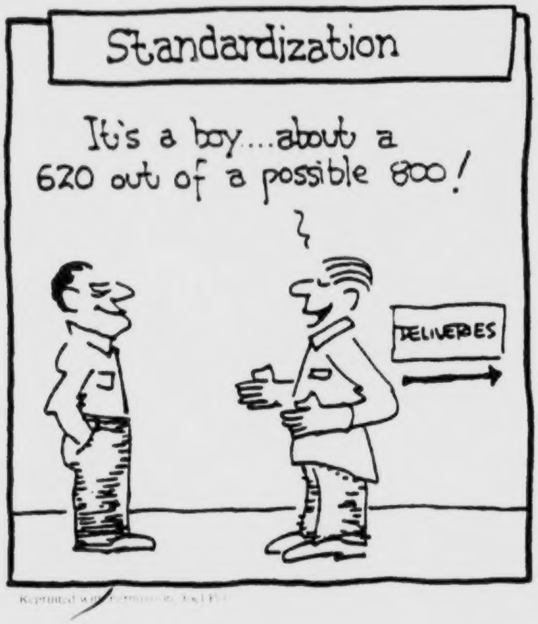
The writing test will assess students in narrative, informative and persuasive writing. Some students will be given extended time periods in order to assess their ability to write longer pieces. Three scoring methods—primary trait, holistic and mechanics—will be used by trained readers. A Writing Portfolio Special Study will be conducted in grade eight.

Although the 1992 reading assessment will include 40% open-ended questions, this is less than the 60% recommended by NAGB's own consensus-building process, which at one point suggested 100% performance-based. In hearings across the nation, witnesses strongly urged NAEP to utilize primarily open-ended reading items. Efficiency and cost were the excuses offered by NAGB in making narrower changes in NAEP than its committee recommended. Critics have charged that NAGB is dragging its feet in developing the new types of assessment needed to assist comprehensive educational reform.

The 1990 math test was described by NAGB as about 20% performance-based. However, Paul LeMahieu, Pittsburgh's testing director, told the National Association of Test Directors that his examination of the test showed that nearly all of the supposed performance-based items were merely multiple-choice type questions without the answer options. As a result, these items were no more capable of assessing higher order or complex thinking than normal multiple-choice questions. LeMahieu warned that the "open-ended" aspects of the reading test may not be any more open ended than were the math items.

Although the Educational Testing Service, the manufacturer of NAEP, has published public relations material touting NAEP as being on the cutting edge of assessment, several state testing programs are actually more innovative. For example, California recently sampled state-wide performance-based math items, and Connecticut is developing performance-based math and science exams for high school that will be implemented in the 1991-92 school year.

Additionally, some NAEP testing may be redundant. Ed Roeber of the Michigan Assessment of Educational Progress, for example, has urged NAGB to coordinate with state assessments to minimize the duplication of effort in developing new assessments and the waste of instructional time taken by testing.


Standardization
It's a boy...about a 620 out of a possible 800!
DELIVERIES

Reprinted with permission [illegible]

K-12 Testing

National Campaign Grows

The Campaign for Genuine Accountability in Education, launched by FairTest and a number of other national organizations with a successful January press conference in Washington, DC, continues to grow. Among the new endorsers are the National Education Association, the Southern Christian Leadership Conference, the Organization of Chinese American Women, National Congress of American Indians, NOW Legal Defense and Education Fund, Advocates for Children of New York, Foxfire, and a growing list of individuals.

The "Statement on Genuine Accountability" signed by the groups called on President Bush and the Governors to avoid reliance on standardized, multiple-choice tests as indicators of educational progress. It urged these leaders to develop new methods of assessment that are educationally sound in order to assess progress toward national educational goals (see *Examiner*, Winter 1990).

At their February meeting, the governors did endorse developing new methods of assessment. Though both the governors and the President sought an annual assessment of students in grades four, eight, and twelve, neither called for using existing multiple-choice tests.

The Campaign will continue efforts to ensure that educational reform is not evaluated with inappropriate tools and that new high quality assessment methods are introduced.

For a copy of the *Statement of the Campaign for Genuine Accountability in Education* and a current list of signers, contact FairTest.

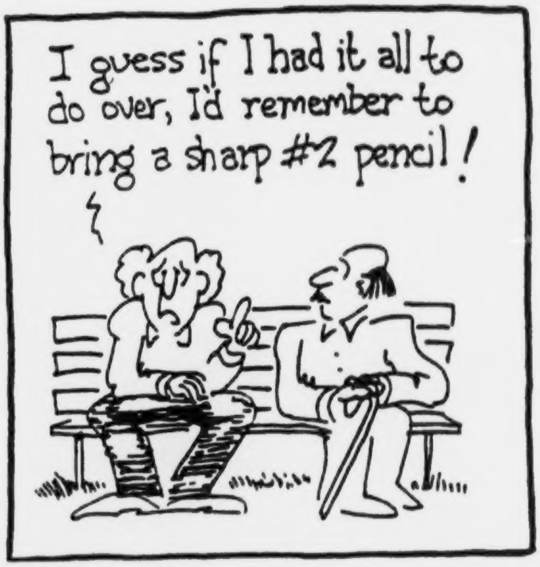
Who Will Control National Assessment?

At its February meeting in Washington, D.C., the National Governors Association (NGA) issued its educational goals for the nation's schools. Broad goals were set for school readiness, school completion, student achievement, math and science, adult literacy and safe schools. Between now and its summer meeting, the NGA will work on developing strategies to attain the goals and measure progress toward them.

The NGA statement indicated that the governors and President Bush agreed to work with the Congress and others to "oversee the development of new measures and indicators" for assessing educational progress. However, a potential battle looms over who will control the oversight process. The president prefers to appoint a body responsible to his administration. The governors, on the other hand, proposed a bipartisan, blue-ribbon group appointed by the President, the governors and the Congress.

Congress, in the Hawkins-Stafford Act, last year mandated the Special Study Panel on Education Indicators to advise the National Center for Education Statistics in the Department of Education. The panel is to report on the quality of policy-relevant education data and recommend improvements in the indicators. The panel's three study groups includes one on readiness for school and educational equity, and another on quality of schools and acquisition of subject matter. Both of these study groups are investigating the role of testing and

(Continued on page 13)


I guess if I had it all to do over, I'd remember to bring a sharp #2 pencil!

Reprinted with Permission Joel Pett

K-12 Testing

NAEP Expansion

(Continued from page 11)

state-by-state comparisons, but the states must pay part of the cost to obtain state data.

NAGB also has proposed speeding up the full implementation of state-by-state comparisons so that, by 1994, NAEP would assess all three grade levels in all major subjects. This would require changing the Act that re-authorized NAEP. Additionally, NAGB proposed setting national educational standards (which they termed "goals") based on responses to selected items on the NAEP exams (see *Examiner*, Winter 1990).

Expansion Too Rapid

Charging that NAGB's proposal for district and school comparisons circumvented agreed-upon procedures for planning any expansion of NAEP, Paul LeMahieu, Pittsburgh's test director and immediate past president of the National Association of Test Directors (NATD), ordered his district out of the 1990 math assessment. He claimed the NAGB plan favored district comparisons even before state comparisons had been tried and evaluated. In a letter to the NATD, LeMahieu accused NAGB of "a serious breach of the faith and trust that we had placed in the process of trial and evaluation *en route* to determining the best possible course for national assessment." LeMahieu charged NAGB with "subvert(ing) the intentions of Congress."

Another critic is Representative Matthew Martinez (D-CA), who described the NAEP expansion as "methodologically unsound" and said it would "neither foster reform nor increase accountability." While the current, limited NAEP has a useful role to play in U.S. education, a greatly expanded NAEP is more likely to stifle education reform than enhance it, Martinez argued in a March 14 *Education Week* article.

Similar concerns were raised by the National PTA in testimony to NAGB in January. Arlene Zielke of the PTA described the proposal and the plan to set national standards based on NAEP test scores as "a power grab without appropriate administrative and policy checks as Congress has intended." The PTA said, "We caution against the creation of a quasi-legislative Board which will replace the Congress as the chief national testing policy maker and pave the way for the erection of a testing empire unresponsive to the needs of parents, teachers, and the community."

FairTest joined with other groups in urging NAGB to slow down its rush to greatly expand testing and instead to concentrate on increasing the use of performance-based items and carefully assessing its own development.

Given limited resources, NAGB will have to choose whether to focus on expansion or improvement. Whether NAEP will now expand depends in large part on 1991 appropriations currently being discussed in the House Appropriations Subcommittee on Labor, Health and Human Services and Education.

Critics have also charged that the 1990 and 1992 assessments have not done enough to introduce cutting-edge assessment methods and that setting standards based on current tests could have harmful effects on the pace and scope of needed educational reforms.

The growing controversy has affected NAGB. At its March meeting, NAGB voted to delay setting standards on the math exams and instead adopted a timetable calling for approval of the framework and definitions for grade-level achievement goals at its May meeting and final action on math goals in November.

A History of Controversy

In renewing NAEP's authorization in 1988, Congress permitted state-by-state comparisons on a voluntary basis. Such comparisons previously had been barred. The legislation established NAGB as an independent governing board and authorized it to set "achievement goals" for NAEP tests. NAGB was also authorized to make recommendations to Congress about changing or expanding NAEP, and charged with developing consensus on the content and form of testing. The tests themselves are produced by Educational Testing Service (ETS) under contract to the department.

Chester Finn, NAGB's chairman, defended the proposed changes as consistent with its role in making recommendations to Congress. The expansion, an increase in test frequency, and a faster schedule for implementing state-by-state comparisons were all proposed in a November 29, 1989 staff paper and approved on December 9 without any public review or input from NAGB's own advisory groups. LeMahieu and others criticized this process as contrary to NAEP's statutory requirement to develop consensus on its future.

(Continued opposite)

K-12 Testing

NAEP Expansion

The staff paper declared the upcoming trial state-by-state math assessment a success due to its endorsement by math educators. But the actual assessment was not conducted until after the paper was approved by NAGB and has not yet been analyzed. A contract to evaluate state-by-state comparisons has been awarded to the National Academy of Education, but another contract, to study the validity of NAEP tests, has not even been awarded. Funds for the validity study were shifted to allow NAEP to overspend its budget. Both studies were authorized to help Congress decide on the future of NAEP after the 1992 reading assessment is completed.

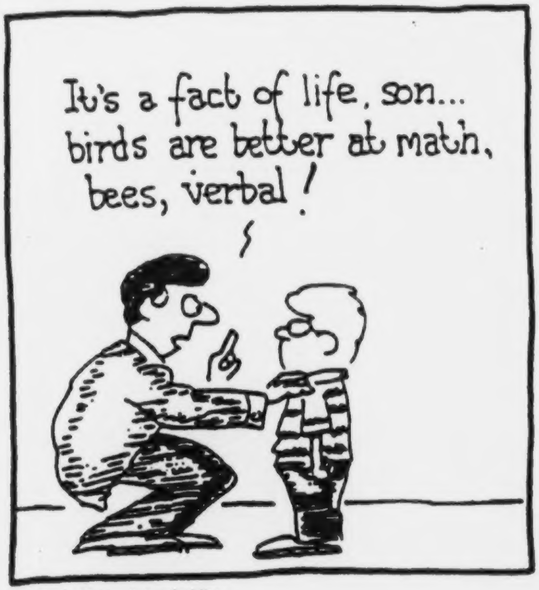
A National Curriculum?

The 1988 NAEP renewal legislation allowed a careful, step-by-step expansion with several trials and studies to ensure adequate test quality and control of unwanted side-effects before full-fledged state comparisons were to be allowed. The current law forecloses complete subject area comparisons until late in the 1990s The NAGB now declares that this timetable is too slow.

The cautious approach chosen by Congress in 1988 was due in large part to concern that NAEP testing would lead to imposition of a national curriculum. Critics of expansion, including FairTest, the National PTA and the American Association of School Administrators, argued that states would encourage teaching to the test in order to score higher than other states, a process that would create a national curriculum. The groups expressed concern that teaching on predominantly multiple-choice and short-response exams would have destructive effects on education. Setting achievement standards based on test scores could have similar controlling effects on curriculum, particularly when combined with state comparisons.

The National Governors Association, in its discussion of assessment, called first for determining what students need to know before defining how and what to assess. NAGB, however, proposed increased testing and setting standards without discussing curriculum. Gordon Cawelti of the Association for Supervision and Curriculum Development (ASCD) told a January 1990 NAGB hearing, "We have got to get the curriculum in order before we set a decent standard." Paul LeMahieu has also criticized what he termed the lack of "educational thinking" underlying the expansion proposal.

At the January hearing, many educators argued that NAGB was not the proper group to set national standards, which should be set through a broader-based process involving at least the governors, Congress and the Bush administration. Educators also renewed their concern that the heavily multiple-choice tests NAEP employed should never be used for setting achievement goals. Local test-score comparisons would create even more pressure for a *de facto* national curriculum defined by a small board and shaped by the content and form of its tests.


It's a fact of life, son... birds are better at math, bees, verbal!

Reprinted with Permission, Joel Pett

University Testing

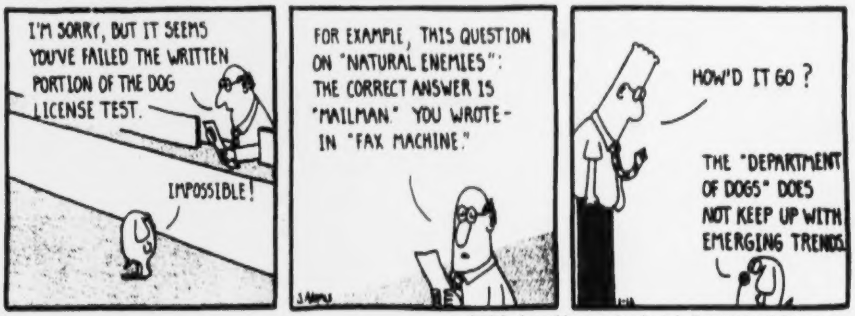
Scores Optional For Admission: 104 Schools

- Antioch College of Antioch University: Yellow Springs, OH *
- Armstrong University: Berkeley, CA
- Bard College: Annandale on Hudson, NY *
- Blue Mountain College: Blue Mountain, MS *
- Boricua College: New York, NY
- Bowdoin College: Brunswick, ME
- Bradford College: Bradford, MA
- Burlington College: Burlington, VT
- **California State University System**
    - California Polytechnic State University, Pomona *
    - California Polytechnic State University, San Luis Obispo *
    - CSU Bakersfield *
    - CSU Chico *
    - CSU Dominguez Hills: Carson *
    - CSU Fresno *
    - CSU Fullerton *
    - CSU Hayward *
    - CSU Long Beach *
    - CSU Los Angeles *
    - CSU Northridge *
    - CSU Sacramento *
    - CSU San Bernardino *
    - CSU San Marcos *
    - CSU Stanislaus: Turlock *
    - Humboldt State University: Arcata *
    - San Diego State University *
    - San Francisco State University *
    - San Jose State University *
    - Sonoma State University: Rohnert Park *
- Central Michigan Univ.: Mount Pleasant +
- Central New England Coll.: Worcester, MA
- **City University of New York (CUNY)**
    - Baruch College *
    - Brooklyn College *
    - City College *
    - College of Staten Island
    - Herbert H. Lehman College
    - Hunter College *
    - John Jay College of Criminal Justice
    - Medgar Evers College
    - York College
- Cleary College: Ypsilanti, MI *
- Cleveland State University: Cleveland, OH *
- Cogswell College: Cupertino, CA *
- College of the Atlantic: Bar Harbor, ME *
- College of Great Falls: Great Falls, MT *
- Columbia College: Chicago, IL *
- Davis and Elkins College: Elkins, WV *
- Dickinson State University: Dickinson, ND *
- Ferris State College: Big Rapids, MI +
- Florida Memorial College: Miami, FL
- Fort Lauderdale Coll.: Fort Lauderdale, FL *
- Franklin University: Columbus, OH *
- Gallaudet University: Washington, DC
- Gannon University: Erie, PA
- Goddard College: Plainfield, VT *
- Golden Beacon College: Wilmington, DE *
- Golden Gate University: San Francisco, CA *
- Governors State Univ.: University Park, IL
- Gratz College: Philadelphia, PA *
- Griffin College: Seattle, WA
- Hampshire College: Amherst, MA *
- Hawaii Pacific College: Honolulu, HI *
- Hendrix College: Conway, AR *
- JFK University: Orinda, CA
- Jones College: Jacksonville, FL
- The Julliard School: New York, NY
- Maryhurst College: Maryhurst, OR
- Medaille College: Buffalo, NY
- Minot State University: Minot, ND *
- Morris College: Sumter, SC *
- National University: San Diego, CA
- **Oregon State System of Higher Education**
    - Eastern Oregon State Coll.: La Grande+
    - Oregon Institute of Technology: Klamath Falls +
    - Oregon State University: Corvallis +
    - Portland State University: Portland +
    - Southern Oregon State Coll.: Ashland +
    - University of Oregon: Eugene +
    - Western Oregon State College: Monmouth +
- Roger Williams College: Bristol, RI *
- Sangamon State University: Springfield, IL
- Shaw University: Raleigh, NC *
- Sierra Nevada College: Incline Village, NV *
- Southeastern University: Washington, DC *
- Southern Utah State Coll.: Cedar City, UT *
- Southern Vermont College: Bennington, VT *
- St. Johns College: Annapolis, MD
- St. Johns College: Santa Fe, NM *
- State University New York College of Technology at Utica/Rome: Utica, NY
- Strayer College: Washington, DC
- Tampa College: Tampa, FL
- Tarkio College: Tarkio, MO *
- Unity College: Unity, ME *
- University of Baltimore: Baltimore, MD
- University of the District of Columbia
- University of Guam: Mangilao *
- University of Houston-Clear Lake: Houston
- University of Texas Health Science Center: Houston
- University of Texas of the Permian Basin: Odessa
- University of Wisconsin-Green Bay +
- VanderCook College of Music: Chicago, IL *
- Villa Julie College: Stevenson, MD *
- Viterbo College: LaCrosse, WI
- Wesley College: Dover, DE *
- West Oahu College: Pearl City, HI
- Wilberforce University: Wilberforce, OH +

*** = SAT/ACT recommended**
+ = SAT/ACT required for class level placement only, not admissions

Information from *Barron's Profiles of American Colleges, 16th Edition, The Right College, 1989* (Arco) and telephone interviews.

Dilbert, by Scott Adams


I'M SORRY, BUT IT SEEMS YOU'VE FAILED THE WRITTEN PORTION OF THE DOG LICENSE TEST.
IMPOSSIBLE!
FOR EXAMPLE, THIS QUESTION ON "NATURAL ENEMIES": THE CORRECT ANSWER IS "MAILMAN." YOU WROTE-IN "FAX MACHINE."
S.ADAMS
HOW'D IT GO?
THE "DEPARTMENT OF DOGS" DOES NOT KEEP UP WITH EMERGING TRENDS.

1990 by United Features Syndicate. Reprinted with Permission

University Testing

More Colleges Make SAT Optional

A recent FairTest survey identified more than 100 colleges around the nation which do not require SAT or ACT scores in their undergraduate admissions process. The institutions range from Bowdoin College in Maine, to the California State University System, to the Oregon State System of Higher Education. Bowdoin, a highly selective private college which made the SAT optional more than 20 years ago, has been very satisfied with the diversity and quality of students it attracts as a result. Both the Oregon State System, which serves 62,000 students on eight campuses, and the California State University System, which serves 355,000 students on 20 campuses, accept students primarily on the basis of a minimum high school grade point average (HSGPA) and high school core curriculum requirements.

Some of these institutions do, however, require SAT or ACT scores for placement or data collection purposes after students are admitted. Using the tests for class assignment forces low-scoring students (who are disproportionately females and members of minority groups) to take and pay for courses that may be unnecessary, despite the fact that neither test has been validated as a placement device. Requiring tests for data collection purposes is unfair to students, who must take the time, spend the money and suffer the anxiety of taking the tests so that schools can receive biographical information free of charge.

Oregon State: GPA Works Best

Concentrating on high school grades as opposed to test scores has produced better students for colleges and universities in the Oregon State System of Higher Education (OSSHE). Recently, OSSHE raised academic standards without focusing on test scores by requiring a minimum high school grade point average, which differs from college to college within the system, and completion of a core high school curriculum.

Data from OSSHE, which enrolls more than 9,000 first-year students annually, show that the system selects a better pool of students by choosing students with the highest HSGPAs rather than those with the highest test scores. In 1988-89, students admitted under the regular requirements (which require a HSGPA minimum but no test scores) outperformed those admitted on the basis of test scores or those admitted on the basis of a "predicted GPA" (an index of test scores and HSGPA). If the Oregon system relied heavily on test scores instead of high school grade point average, it would admit a significantly less academically-able student body (see chart below).

Despite having slightly lower SAT-Math scores, those admitted on the basis of a 3.0 minimum HSGPA earned significantly better grades in all levels of math courses than those admitted on the basis of test scores. The same disparity holds true for the SAT-Verbal section.

These data contradict one favorite myth of the College Board and ETS: that test scores are a universal measure or "common yardstick" that enable schools to make better admissions

(Continued on page 14)

Student Achievement by Admissions Alternatives

Students admitted to the Oregon State System on the basis of:

	High School GPA	**"Predicted GPA" (HSGPA and test scores)**	**SAT/ACT Scores**
SAT average:	947	972	986
High School GPA received:	3.29	2.72	2.31
First Year College GPA received:	2.68	2.36	2.22

(source: Office of Academic Affairs, Oregon State System of Higher Education, 12/21/89)

University Testing

News in Brief

CA Summit Rejects SAT

The final report of the California Education Summit, which was held last December and attended by superintendents, university faculty and administrators, has recommended that the state's two higher education systems deemphasize the SAT in admissions and placement. The report states that "California's higher education systems [University of California and California State] should be encouraged to accept end-of-course exams for admission and placement purposes in lieu of Scholastic Aptitude Test (SAT) results and examinations developed and used by those systems." The University of California system currently requires all applicants to submit either SAT or ACT scores. The California State University system does not require any test scores for admission (see story, p.9), but students may submit them if they do not meet a minimum high school grade point average. Some colleges in the system may also use the tests for placement of students in remedial or upper level classes.

NCAA Regulations Harm African Americans

In 1989, as in past years, a large majority (66%) of student athletes denied freshman eligibility because they meet some, but not all, of the National Collegiate Athletic Association's (NCAA) "academic standards" are African American. (Under the NCAA's Proposition 48, first year students who do not meet an SAT or ACT test score cut-off and a minimum high school grade point average in core subjects are barred from interscholastic competition.) This result is not surprising: 75% of students who fall into this "partial qualifier" category do so because of the test score cut-off rather than the GPA minimum. Since African Americans average more than 200 points lower than whites on the SAT, due, in part, to the test's unfair biases, they will be greatly overrepresented among those students barred from playing. In fact, 16% of all African American athletes admitted to Division I institutions were deemed partial qualifiers, compared to 2.5% of white athletes.

Same Scores, Higher Grades For Female Athletes

Another recent study by the NCAA shows that, despite similar SAT scores, female basketball players at Division I schools earn higher grades than male basketball players in both high school and college, a finding consistent with the performance of men and women on the SAT and ACT overall.

Mass. Legislative Committee Endorses SAT-Optional

The Massachusetts legislature's Education Committee has approved a resolution calling on the state university system to make submission of SAT scores optional in its admission process.

The resolution, filed by Representative Barbara Gardner, "strongly encourage(s) the Board of Regents of the Commonwealth of Massachusetts to discontinue the use of standardized test scores as a required part in the process of determining eligibility for attendance at a campus maintained by the state . . ." Last year, Massachusetts ended funding for its Senatorial Honors Scholarships, which were awarded solely based on SAT scores, after FairTest documented that 60% of the awards went to boys, due to the test's gender bias.

The next step for the SAT-optional resolution is the full House, which is expected to take up the proposal later this session.

Harvard Law Students Call for Changes in Admissions Process

The Harvard University Law School Council, the student governing body, recently issued a white paper calling for deemphasizing traditional critieria for admitting students, including scores from the Law School Admissions Test (LSAT).

The report, *A Program In The Public Interest For Harvard Law School*, states, "In addition to admitting candidates with demonstrated and proclaimed commitment to public service, Harvard Law School must also admit candidates who demonstrate that they possess the highest of analytical, reasoning and writing skills by means other than grades and LSATs. This can be done by reference to published written reports, speeches and other evidence of a candidate's initiative, creativity, and leadership abilities. Greater emphasis must be placed on such non-traditional criteria in recognition of the often superior demonstration of abilities they reflect, particularly for older students."

For a copy of the white paper, write to Law School Council, Harvard Law School, Cambridge, MA 02138.

University Testing

ETS/CEEB Lawsuit

(From page 1)

York passed Truth-in-Testing in 1979 ,the test manufacturers cut back on the number of times tests were given to punish the state for trying to protect its consumers. Now they say they are doing us a favor by offering more test administrations in New York at the expense of students' rights."

In response to the legal action, Ralph Nader stated, "ETS, the College Board and their lawyers have made a big mistake. This lawsuit will make the standardized testing industry fraud a nationally visible issue for public discussion."

"ETS President Greg Anrig's position is monumental hypocrisy," added FairTest Public Education Director Bob Schaeffer, co-author of *Standing Up to the SAT*. "When Greg Anrig was Massachusetts Commissioner of Education, he endorsed Truth-in-Testing and told the *Boston Globe*, 'Decisions that affect a person's career or life can be made by these tests. Such matters should be subject to public scrutiny.'"

Apparently Anrig, now the President of the Educational Testing Service, changed his tune after a recent federal court ruling found that Truth-in-Testing conflicted with federal copyright law. However, that decision applied only to the Medical College Admissions Test (MCAT) and is currently being appealed by the state (see related story at right).

FairTest and other defenders of the free flow of test information, including Senator Ken Lavalle, the author of Truth-in-Testing, are considering a variety of measures to preserve the law.

Truth-in-Testing Appealed

The State of New York has appealed a controversial federal district court decision striking down New York's ten-year-old Truth-in-Testing law. In January, Judge Neil P. McCurn of the Northern District Court ruled that the state's Truth-in-Testing provisions are preempted by federal copyright law.

The Truth-in-Testing law requires testing companies to publicly disclose university admissions test questions and answers, as well as make research reports on test validity and fairness available to the public.

While the decision striking down the law applies only to the plaintiffs in the case—the American Association of Medical Colleges, producers of the Medical College Admissions Test (MCAT)—it leaves the door wide open for challenges such as that made by the Educational Testing Service and the College Board to exempt some of their tests from the law (see related story, page 1).

In a letter to the Solicitor General of New York urging the state to appeal the decision, FairTest explained the importance of the law: "Until the legislature passed the Truth-in-Testing Act in 1979, the testing industry was completely unaccountable to the citizens of New York. After the law went into effect, this situation changed dramatically. First, students were able to challenge test questions and answers. As a result, thousands of students' scores were raised, and the entire testing experience became a more fair and open process.

"Second, the measure enabled independent organizations and researchers to monitor the testing industry and critique its products. Third, the testing companies, aware that legislation modeled after New York's law was pending in nearly 20 states and in the U.S. Congress, voluntarily extended Truth-in-Testing throughout the country. In short, the Truth-in-Testing Act has had tremendous significance, not just in New York, but nationwide."

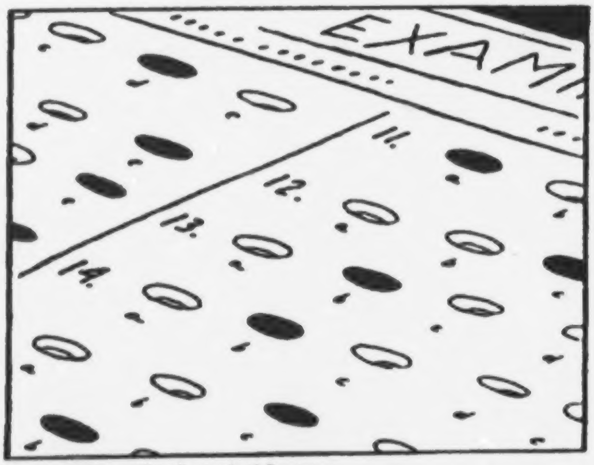


Illustration by Michael Klein. Reprinted with Permission

K-12 Testing

Grade School Testing in Brief

SACUS Seeks Ban

The Southern Association on Children Under Six (SACUS) has joined the growing chorus of groups calling for a ban on testing young children. In a newly adopted position paper on "Developmentally Appropriate Assessment," SACUS "recommends a ban on the routine, mass use of standardized intelligence, achievement, readiness and developmental screening tests for young children through the age of eight." The paper briefly discusses how testing practices harm children, the limitations of standardized testing, misuses of test data, appropriate assessment, criteria for good assessment, and good assessment practices. Copies are available from SACUS, P.O. Box 5403 Brady Stn., Little Rock, AR 72215; $.30/ea. ($.20 ea for members), 10% discount on 10 or more.

NY Campaign Gets Underway

A coalition of groups, organized by the New York Public Interest Research Group (NYPIRG), the American Reading Council and FairTest, has formed the *New York Campaign for Genuine Accountability in Education*. As its first public act, the Campaign has issued a "Statement to Urge a Ban on Standardized Testing of Young Children." The statement calls on the state's Board of Regents and local school boards to "ban the mass use of standardized tests from pre-school through second grade" and replace these tests with developmentally appropriate evaluation procedures. The proposed ban would exempt bilingual education. To date, the Statement has also been endorsed by the Public Education Association, the United Parents Association, the N.Y. State Child Care Coordinating Council, the NAACP Legal Defense and Educational Fund, and the New York State Association for the Education of Young Children. Many other organizations are now considering it.

To expand its work, the Campaign has begun statewide education and training sessions on the issue. FairTest and NYPIRG have received a grant from the Aaron Diamond Foundation to conduct this work statewide. For more information, contact the Campaign c/o Joan Harris, NYPIRG, 9 Murray St., NY, NY 10007-2272; (212) 349-6460.

Kentucky Plans Overhaul of Assessments

As part of a comprehensive overhaul of its educational system, Kentucky plans to develop new methods of assessing student achievement. Achievement will be one of the factors included in a new system of rewards and sanctions based on school performance. The state restructured education following a Kentucky Supreme Court order to revise school financing and other areas of the system.

KY Rankings Misleading

University of Kentucky researchers Thomas Guskey and Edward Kifer have concluded that Kentucky's process of ranking schools by statewide test data "can be problematic and potentially misleading." Writing in *Educational Measurement* (Spring 1990), they report that the rankings "do not even serve their intended purpose, which is to tell ... how well or poorly schools or school districts are doing." They found that "nearly 42% of the variance in the total achievement scores could be explained" by the percent of economically deprived students in a district. District rankings were heavily affected by the choice of data used to determine the ranking. Using different subsets of data (e.g., reading or math scores) produced different rankings, and the correlations among rankings were low. Thus, for example, a high overall score could disguise weakness in some areas, or a low overall score could hide strengths in some areas.

Illustration by Michael Klein. Reprinted with Permission

Honig Praises New Writing Assessment in CA

Though no improvement was demonstrated on California's performance-based writing test over the two previous administrations, Superintendent Bill Honig praised the program

(Continued opposite)

K-12 Testing

Grade School Testing in Brief

(From left)

for having a positive effect on classroom activity. He told the *Los Angeles Times*, "To show continued success on the CAP writing test, schools must develop a multi-year approach to writing instruction that crosses all grade levels and subject areas." He added it was California's "first major step to move beyond multiple-choice tests." A state education summit in December called for replacing all multiple-choice testing with performance-based tests over the next five years (see Quote of the Month on page 2).

Eastern European Upheaval Changes Test

North Carolina has been forced to change two items on its sixth-grade social studies test to keep up with events in Eastern Europe. The "wanted" answer to one item said that in the Soviet Union, citizens could only vote for "candidates selected by the Communist Party." The second item asked, "Which of these countries has a democratic government? A: Albania, B: Cuba, C: United States, D: Poland." Since voters in the Soviet Union's most recent election could choose among competing candidates and Poland is now classified as a democracy, more than one answer is correct on each item. Initially, testing officials instructed teachers to tell their students to answer the questions as if Eastern Europe had not changed since 1987. After that directive became public, the Department of Public Instruction decided to re-write the items.

Control of National Assessment

(From page 5)

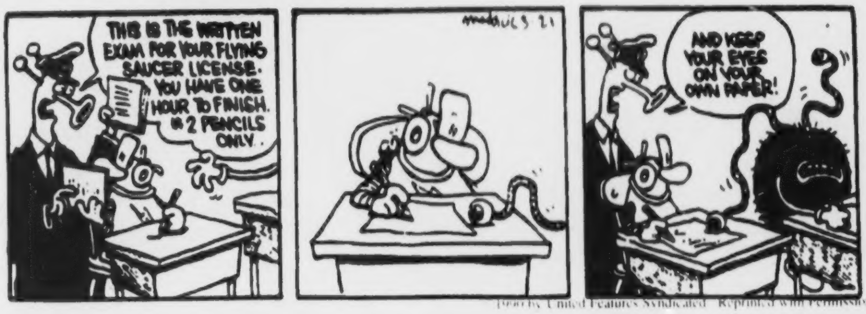
assessment. The six areas the panel is investigating overlap substantially with the governors' and administration's six goals.

Meanwhile, Sen. Jeff Bingaman (D-NM) has introduced legislation to set up a National Council on Educational Goals to analyze existing information on U.S. schools and to develop a "comprehensive set of national goals that will help ensure excellence in education." Bingaman's Council would issue annual "report cards" and make recommendations for improving the assessment of education and the collection of education information. Initial appointment of the nine voting members of the Council would be made by the president, with approval from the National Governors Association and consent of the Senate. The Council would then be self-perpetuating. The Chair of the National Assessment Governing Board (NAGB) would be one of three ex-officio members of the Council, but the power of NAGB to set the nation's assessment agenda could be sharply reduced by the creation of a Council with a broader mandate than NAGB possesses. NAGB was set up only two years ago as an independent body to oversee the development of NAEP.

Young Children

Another part of the governors' statement reads, "By the year 2000, all children in America will start school ready to learn." Kids, Education and You (KEY), a parent group from Denver City, TX, raised concern that "ready to learn" will provide an opening for national expansion of "readiness testing" such as the much-criticized Gesell instruments. But Roger Porter, President Bush's domestic policy advisor, wrote FairTest and KEY that the President does not intend "readiness" to mean readiness testing. The governors' statement does not address this issue, but does say, "Placement decisions for young children should not be made primarily on the basis of standardized tests," and that schools "must be able to effectively educate students" regardless of variations among the students.

Robotman, by Jim Meddick


THIS IS THE WRITTEN EXAM FOR YOUR FLYING SAUCER LICENSE. YOU HAVE ONE HOUR TO FINISH. #2 PENCILS ONLY.
Meddick 2-21
AND KEEP YOUR EYES ON YOUR OWN PAPER!

[illegible] by United Features Syndicated. Reprinted with Permission

BEST COPY AVAILABLE

SAT Optional

(From page 9)

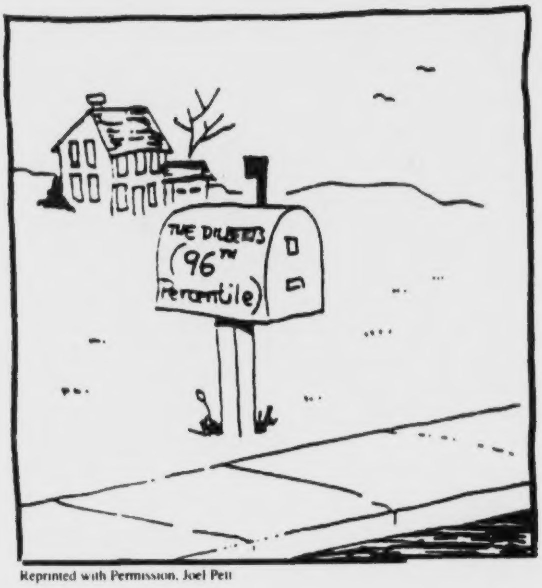
decisions. At least at Oregon's state universities, students who would appear to be stronger on the basis of their test scores are, in fact, lower-performing students.

Cal. State: Quality in Diversity

The California State University system's mission includes providing access to higher education for traditionally excluded groups. By de-emphasizing test scores, CSU has simultaneously succeeded in raising academic standards and promoting diversity: the student body in CSU institutions is now 35 per cent minority and 54 per cent female. More than two-thirds of CSU first-year students are admitted solely on the basis of a minimum HSGPA of 3.0 and a core high school curriculum. Only students who fail to achieve the GPA standard and wish to be considered on the basis of a test score/HSGPA index are required to submit standardized test scores.

CSU leaders are fully aware of the connection between test scores, lack of diversity and poor educational quality. One official told FairTest, "I don't like test scores, and all the top administrators feel the same way. In fact, the whole CSU system has a mistrust of multiple-choice instruments. We've raised standards without raising test scores by focusing on core curriculum and HSGPA minimum." Pointing out the correlation between test scores and a student's socio-economic status, he said, "We know from experience that if a kid meets the HSGPA minimum but their test scores are low, that usually means that their parents didn't go to college."

Administrators in the CSU system are also resisting the use of multiple-choice tests in measuring "basic skills" and "outcomes" of higher education. For that reason, they are considering, among other things, a portfolio model of assessment rather than test scores. Due to resource constraints, any assessment instrument must be part of the regular curriculum. But that is the most appropriate and authentic approach to assessment in the eyes of many educators.



Reprinted with Permission, Joel Pett

Bar Exam

(From page 16)

York State Judicial Commission on Minorities, which is studying New York's bar examination process, said that only about a third of African Americans who take the exam pass. Former New York City Corporation Counsel Peter Zimroth concurred, saying that as many as two-thirds of the minorities working in his office failed the exam the first time, despite satisfactory performance as lawyers. Stanley Chess, President of BAR/BRI Bar Review, called for an overhaul of New York's bar examination, particularly the multiple-choice section.

Participants also noted that while law school curricula have changed dramatically to include clinical education, the bar examination has failed to reflect this trend. Some suggested that a one-year apprenticeship program with a practicing lawyer would be a more meaningful gatekeeper to the profession than the current system. Others believed that graduation from an accredited law school that included clinical training should be adequate preparation to practice.

These and other proposals are currently being studied by the Legal Education and Bar Admission Committee of the New York City Bar Association, which plans to release a report in the next year with recommendations for changes in the bar admission process.

FairTest Publications

_______ ***FairTest Examiner***, published quarterly. . . $15.00/year (individuals), $25.00 (institutions)

_______ Back Issues of ***FairTest Examiner***, $4.00 each: Vol. ___No. ___(All 11 for $32.00)

_______ ***Standing Up to The SAT***, by John Weiss, Barbara Beckwith and Bob Schaeffer (192 pp.) . . . $6.95

_______ ***Fallout From the Testing Explosion: How 100 Million Standardized Exams Undermine Equity and Excellence in America's Public Schools***, by Noe Medina and Monty Neill (3rd ed., 77pp.) . . $8.95

_______ ***Sex Bias in College Admissions Tests: Why Women Lose Out***, by Phyllis Rosser and FairTest staff, 3rd ed. (60 pp.) $7.95

_______ ***The Reign of ETS***, by Allan Nairn and Ralph Nader (550 pp.). . . $30

_______ ***Beyond Standardized Tests: Admissions Alternatives That Work***, by Amy Allina with FairTest staff (18 pp.) $5.50

_______ ***None of the Above, Behind the Myth of Scholastic Aptitude***, by David Owen (327 pp.) . . . $5.95

Contributions to FairTest

I agree there's no better way to be part of the testing reform movement than supporting FairTest! Enclosed is my tax-deductible contribution!*

_______ **FairTest Associate** ($25.00--includes a free one-year subscription to the *FairTest Examiner*)

_______ **FairTest Sponsor** ($50.00--includes *Examiner* subscription and *Standing Up to the SAT*)

_______ **FairTest Sustainer** ($100.00 or more--includes above plus a 20% discount on publications)

_______ **Other**

Publications Total _______________

Contributions Total _______________

Check Total _______________

Name ____

Address ____

City ____________________ State ______________ ZIP ______________

Please make checks payable to FairTest and mail to: FairTest, 342 Broadway, Cambridge MA 02139

**Contributions to FairTest are fully tax deductible, except for the value of publications received.*

Public Hearing and Articles Spark Discussion

Bar Exam Debated

The bar examination—the test that determines who can and cannot enter the legal profession—is coming under increased scrutiny. Nowhere is this more evident than in New York, where a series of articles, hearings and forums held over the past year have questioned the validity and fairness of the test.

Public Hearing

"It is time for the State legislature to take a look at qualifications for the practice of law in New York," said Assemblyman G. Oliver Koppell. "A substantial number of law school graduates have difficulty passing the bar exam, and questions have been raised as to its relevance and whether, in fact, it measures the skills and knowledge required to become a competent lawyer." Koppell, Chair of the Assembly Committee on Judiciary, presided over a public hearing on the bar on March 15 in New York City.

At the hearing, FairTest provided testimony on the exam's disparate impact on minorities and called on the state to create a fairer, more accountable and professionally-sound method of licensing lawyers. FairTest called for the establishment of a special committee to examine alternatives to the current bar examination system, including a performance-based test that would more closely reflect the day-to-day duties of an attorney. The committee would also examine problems such as the lack of due process rights for test-takers who challenge test questions, the use of bar results to evaluate the quality of law schools, and the Board of Law Examiner's failure to disclose past examination questions so that they may be subject to independent review.

New York State Bar Association

Two articles in the *New York State Bar Journal* have also sparked discussion. In "Racial and Ethnic Barriers to the Legal Profession: The Case Against the Bar Examination" (April 1989), Maurice Emsellem concluded, "That the bar examination has a discriminatory impact on the minority examinee is undisputed."

Emsellem noted that in California, the only state which regularly collects and analyzes this data, 65% of whites received passing scores on the bar exam, while only 9% of African Americans were successful in 1988. Over the past ten years, the pass rate for whites has averaged twice that of minority examinees. Emsellem also criticized the test on the basis of its "scientific shortcomings," most notably the fact that no studies have attempted to correlate bar exam scores with success as a lawyer. In a follow-up article defending the exam, entitled "Bar Examinations: Ignoring the Thermometer Does Not Change the Temperature" (October 1989), Stephen Klein, a consultant to the California Bar Examiners, did not address this fundamental issue of validity.

The bar exam was also the subject of a panel at the annual meeting of the New York State Bar Association in January 1990. At this meeting, several panelists and audience members criticized the bar for excluding qualified minorities from the profession, while not meeting its stated goal of protecting the public from incompetent lawyers.

Edna Wells Handy of the New

(Continued on page 14)

FairTest®
National Center for Fair & Open Testing
342 Broadway
Cambridge, MA 02139-1802
(617) 864-4810
(617) 497-2224 (FAX)

Non-Profit
U.S. Postage
PAID
Cambridge, MA
Permit No. 53278

ADDRESS CORRECTION REQUESTED

Larry Roudner
ERIC/TME
3333 K St. NW #200
Washington, DC 20009-1275

Moving? Return your label with new address

FairTest

EXAMINER

The National Center for Fair & Open Testing Vol. 4, No. 3 Summer, 1990 ISSN # 0898-2511

National Testing Commission Issues Final Recommendations

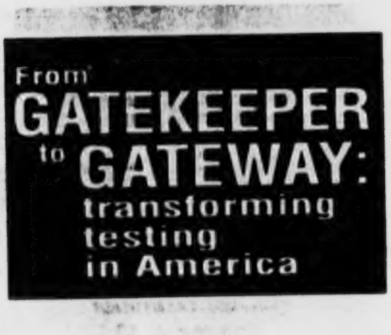
New Report Echoes FairTest Agenda

After an exhaustive three-year study, a blue-ribbon commission has called for the comprehensive restructuring and reduction of educational and employment testing.

From Gatekeeper to Gateway: Transforming Testing in America, the report of the National Commission on Testing and Public Policy, concluded that "current testing, primarily multiple-choice in format, is overrelied upon, lacks adequate public accountability, sometimes leads to unfairness in the allocation of opportunities, and too often undermines vital social policies."

Among the key findings and recommendations of The Ford Foundation-funded Commission are:

- "Testing programs should be redirected from overreliance on multiple-choice tests toward alternative forms of assessment."
- "Students are subjected to too much testing in the nation's schools."
- "Test scores are imperfect measures and should not be used alone to make important decisions about individuals, groups or institutions; in the allocation of opportunities, individuals' past performance and relevant experience must be considered."
- "No test can be wholly free of cultural bias....We must stop pretending that any single standard can illuminate equally well the talents and help promote the learning of people from different backgrounds."



From GATEKEEPER to GATEWAY: transforming testing in America

Report of the National Commission on Testing and Public Policy

- "The more test scores disproportionately deny opportunities to minorities, the greater the need to show that the tests measure characteristics relevant to the opportunities being allocated."
- "The enterprise of testing should be subjected to greater public accountability."

FairTest, which testified at three of the Commission's five public hearings, endorsed the report and called for the implementation of its proposals. In a letter to Commission Executive Director George Madaus, FairTest said that it is "gratified that the in-depth research conducted by this most distinguished Commission has led our institutions to similar conclusions. Agreement with the Commission's conclusions, however, is not enough. FairTest is committed to the implementation of testing reforms to bring these important recommendations to fruition. That is why FairTest will continue to lead a national campaign to make certain that evaluation of students and employees is fair, open, accountable and educationally sound."

In a statement released at the National Press Club news conference unveiling the Commission report, FairTest stated, "As the report makes

(Continued on page 11)

CONTENTS

New FairTest Field Organizer 2
The ACT: Bias & Misuse 3
Writing/Reading Inadequate 6
Call to Halt NAEP Expansion 7
NY Campaign Underway 8
Assessing LEP Students 9
FairTest Order Form 15
Performance-Based Assessment Spreading 16

TM 017 189

FairTest Hires Field Organizer

After an extensive national search, FairTest has hired Cambridge resident Veda Wright as its Testing Reform Field Organizer. Well known in the Boston area for her work with community education organizations, Ms. Wright brings a wealth of experience to her new job with FairTest.

Ms. Wright spent the last two years working for Roxbury Children's Services at Madison Park High School, located in the heart of Boston's African American community. She was also on the staff of the Twentieth Anniversary Mobilization, which coordinated the Massachusetts delegation to the Washington commemoration of Dr. Martin Luther King Jr.'s famous 1963 March. In addition, Ms. Wright has worked with a number of education organizations and has been active in local volunteer programs as a mediator in a local court, a member of the Cambridge Commission on the Status of Women, and a parent activist.

Veda Wright

The staff and board of FairTest welcome Veda, and hope that many *Examiner* readers will have the opportunity to meet her in the near future. ⌝

FairTest
National Center for Fair & Open Testing
342 Broadway
Cambridge, MA 02139-1802
(617) 864-4810
(617) 497-2224 (FAX)

FairTest Examiner
Volume 4, #3

Staff: Mike Friedman, Publications Director/Office Manager; Monty Neill, Associate Director; Bob Schaeffer, Public Education Director; Cinthia Schuman, Executive Director; Jamie Souder, Administrative Assistant; Sarah Stockwell, University Tests Coordinator; Veda Wright, Testing Reform Field Organizer.

Board of Directors: Denise Carty Bennia, Esq., CUNY Law School, Chair; Dr. Pamela George, North Carolina Central University; Dr. Susana Navarro, California Achievement Council; Paul Pottinger, Goodsmith and Co.; William Robinson, Esq., Dean, D.C. School of Law; Bob Schaeffer, Treasurer; Steve Solomon, Esq., Proskauer, Rose, Goetz & Mendelsohn; Chuck Stone, *Philadelphia Daily News*; Ron Vera, Esq., Barbosa & Vera; John Weiss, Public Policy Initiatives, President.

National Policy Panel: Dr. Leonard Beckum, City University of New York; Dr. Beverly Cole, NAACP; Dr. Harold Dent, Association of Black Psychologists; Arnold Fege, National PTA; Dr. Howard Gardner, Harvard University; Dr. Asa Hilliard, Georgia State University; U.S. Rep. Matthew Martinez (D-CA); Deborah Meier, Central Park East Secondary School; Sara Melendez, University of Bridgeport; Aurelio Montemayor, Intercultural Development Research Association; Harry Pachon, National Association of Latino Elected and Appointed Officials; Alfredo de los Santos, Maricopa Community Colleges; Eleanor Smeal, Fund for a Feminist Majority; Theodore Shaw, Esq., NAACP Legal Defense Fund; Dr. Gail Thomas, Texas A & M University; U.S. Rep. Ted Weiss (D-NY).

Funding for FairTest's work comes from the Golden Rule Corporation, the Rockefeller Family Fund, the Ford Foundation, the Phillip M. Stern Family Fund, the Aaron Diamond Foundation, the Edward Hazen Foundation, the Lilly Endowment the New York Community Trust, and many individual contributors.

New FairTest Donors

FairTest is pleased to announce the receipt of several new grants from leading foundations in the field of education. **The Aaron Diamond Foundation, Edward W. Hazen Foundation, Lilly Endowment** and **New York Community Trust** have all made recent contributions to our testing reform work at the kindergarten through grade 12 level.

The **Golden Rule Insurance Company, Ford Foundation, Rockefeller Family Fund** and **Philip M. Stern Family Fund** continue to provide assistance for FairTest's general program work.

FairTest is grateful to all of its foundation funders, as well as those of you who have become FairTest contributors in the past year. Your support is critical to ensuring that evaluation is fair, open, accountable and educationally sound.

For information on how to become a FairTest contributor, see page 15. ⌝

University Testing

The ACT: Bias, Imprecision and Misuse

To many people, the ACT Assessment, made by the American College Testing Program of Iowa City, Iowa, is "the other college admissions test." Compared to its more visible and larger multiple-choice cousin, the SAT, it receives relatively little attention or criticism. Yet the ACT is very influential: more than 850,000 high school students take it annually, and it is the predominant college admissions test used in the Midwestern, Southern, Rocky Mountain and Plains states.

Unfortunately, the ACT shares many of the SAT's flaws. For instance, both tests predict first year college grades less well than high school grades do, and both underpredict the college performance of young women. As with the SAT, the ACT's lack of precision, coupled with its gender, race, income and age bias, make it of questionable value. Pervasive misuse of ACT scores can also do great harm to students.

To improve its position in the college admissions testing marketplace, an "enhanced" version of the ACT Assessment was released last fall. The new test purports to be more responsive to high school curricula and to colleges' expectations about students' academic backgrounds. The changes, according to the test makers, "involve *test content* and *test scores*. The new tests... measure a broader array of knowledge and skills than did the original tests" (emphasis in original). However, "All other components and features of the ACT Assessment remain essentially unchanged." Like the old version, this year's ACT is scored on a scale of 1-36, though the scores for the old and new tests are not directly comparable.

While its manufacturers make many claims about the "Enhanced ACT Assessment," none of the changes made on the new test address the issues of gender, race and other cultural bias. Therefore, studies having to do with unfair bias, such as those addressing the differences between male and female or minority and majority performance on the old ACT, are still relevant.

Cultural bias clearly exists on the ACT. Significant unexplained test score differences remain between minorities and whites, even when students are grouped by socioeconomic status, educational background, high school coursework, and class performance. For example, ACT researchers found that African American seniors average 13.9 compared to 18.3 for whites with the same grades, class rank and socioeconomic status.

A similar study of white and Hispanic performance on the ACT showed that even when Hispanic and white students have taken the same rigorous college preparatory science and math courses (chemistry, trigonometry, and calculus), Hispanic students still score 2.8 points lower on the ACT mathematics test. At every income level whites significantly outscore Hispanics.

(Continued on page 4)

1989 ACT AVERAGE COMPOSITE SCORES BY GENDER AND RACE

(1 to 36 point scale)

Asian M=20.5
White M=20.2
Asian F=19.5
White F=18.8
Puerto M=17.9
Puerto F=16.3
Mexican M=16.2
Indian M=15.2
Mexican F=14.8
Indian F=14.3
Black M=13.8
Black F=13.4

Asian = Asian Americans
Indian = American Indians
Mexican = Mexican Americans/Chicanos
Puerto = Puerto Ricans/Cubans and other Hispanics

(*Source: Special ACT Assessment Reference Norms*, National 1989.)

University Testing

The ACT Assessment—

As with minority groups, the bias against females is evident in unexplained score gaps between men and women. Furthermore, the ACT underestimates the grades of young women. According to the manufacturers, the ACT's sole scientific purpose is to predict the first year college grades of test-takers. Yet the ACT underpredicts grades for girls, while overpredicting boys'. Thus, the ACT indicates that females' college performance will be worse than males' even though females actually earn higher grades than males in both high school and college.

A study co-authored by an ACT researcher found that the "differences in the predictions are of sufficient magnitude to be of practical importance." In fact, "A large proportion of students is generally affected by either over- or underprediction" (Leann M. Gamache, ACT, and Melvin R. Novick, University of Iowa, 1985). Not surprisingly, females score lower than males within each racial/ethnic group, creating a double jeopardy for women who are members of minority groups (see chart, p. 3).

Older students also face discrimination from the ACT. Like females, students who don't go to college right after high school consistently have their college grades underpredicted by the ACT. This may deny them important educational and scholarship opportunities.

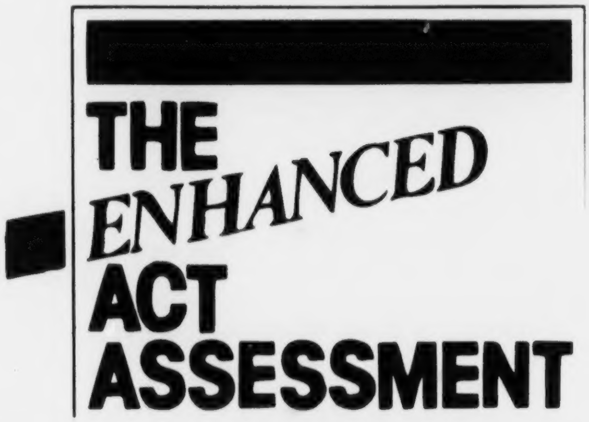
FairTest and other testing reform advocates maintain that bias in the ACT and SAT exaggerates already existing differences between whites and members of minority groups: reliance on college admissions test scores makes an inequitable educational system worse. The National Commission on Testing and Public Policy pointed out some of the sources of bias inherent in all standardized tests in its recent report (see p.1):

There are two broad reasons for group differences in test performance. "First, tests, as well as other methods of predicting behavior, are culturally bound and almost always reflect the dominant or "national" culture in both form and content. Research shows that the way test content is oriented —toward the topics and culture of one group as opposed to another — can significantly affect test scores... The second cause of differential group performance on tests is economic and educational."

Lack of precision in ACT scores makes matters worse for all affected groups. The new ACT has a margin of error of plus or minus one point on the Composite score (on a scale of 1 to 36), plus or minus two points on each test (i.e. English, Math, Reading, and Science Reasoning, on scales of 1 to 36), and plus or minus two points on each subtest (on scales of 1 to 18). This means that a test-taker's "true" score on a subtest can fall anywhere within a range of more than four points out of a total of 18.

This imprecision makes use of ACT cut-off scores (i.e. absolute minimum scores) for admission to college, awarding of scholarships, or academic placement decisions particularly arbitrary and unfair. Those who might benefit from a college education or scholarship the most—members of minority groups and lower-income students—are also those most likely to fall short of a cut-off score by a point or two because of unfairly low scores.

Many groups, such as the National Commission on Testing and Public Policy, strongly criticize the use of cut-off scores, asserting that a candidate's past performance should always



THE ENHANCED ACT ASSESSMENT

While the manufacturers make many claims about the "Enhanced ACT Assessment," none of the changes made on the new test address the issues of gender, race and other cultural bias.

University Testing

Bias, Imprecision and Misuse

(from page 3)

be considered in the allocation of opportunities. Similarly, the guidelines of the National Association of College Admissions Counselors direct members not to use cut-off scores: "College and University Members agree that they will refrain from using minimum test scores as the sole criterion for admission, thereby denying certain students because of small differences in scores."

Despite these cautions and ACT's own test use guidelines, which state that "ACT test score data should not be used as the sole criterion for admission/selection decisions," ACT actually condones and even subtly encourages the use of cut-off scores in its promotional materials. For instance, ACT proposes that colleges and universities use a "Simplified Admissions System," in which "the ACT student record can be used as a 'total' admission system." One "common component" of such a system, according to ACT, is that it is "generally restricted to specific categories of students (e.g. . . . students above a certain ACT Composite score.)" Implementing this amounts to imposing a cut-off score.

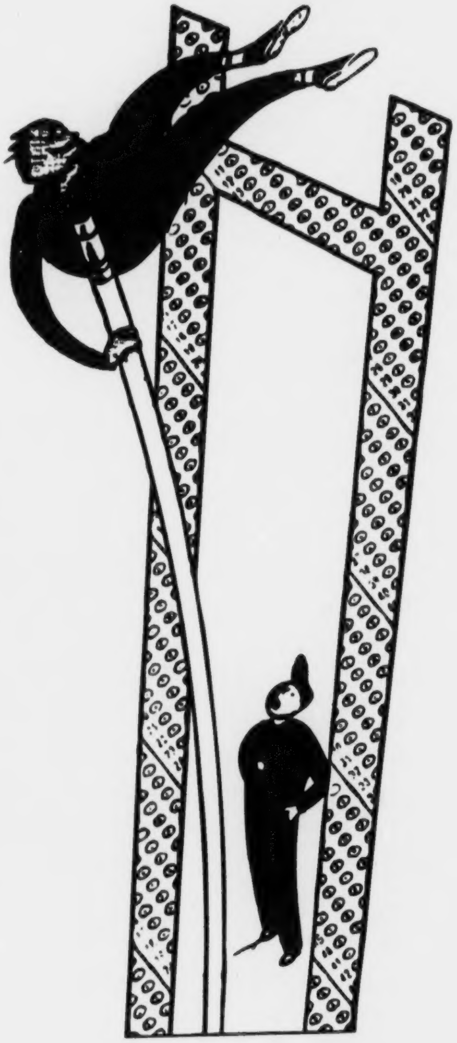


Illustration by Steve McCracken. Reprinted with permission

Such disparity between practice and principle is partly responsible for the pervasive misuse and overuse of the ACT. Colleges and universities frequently list ACT cut-off scores in college guidebooks and the National Collegiate Athletic Association uses an ACT or SAT cut-off to deny athletic eligibility to students who are supposedly academically underqualified. Some institutions, such as the State Board of Education in Alabama, use an ACT cut-off in awarding scholarships. This practice resulted in *all* 36 scholarships for teacher candidates being awarded to whites from 1982 to 1984, despite Alabama's large African American population. ACT could stop reporting test scores to institutions which persist in misusing them, but it has never taken such action.

Across the nation, decisions are routinely being made about peoples' educational opportunities and ultimately their lives on the basis of the ACT, a biased and imprecise instrument. That is why FairTest believes that the ACT, like the SAT, should be made optional in the college admissions process. ❒

NY Lawsuit Article Available

FairTest is offering reprints of a law journal article about the 1989 lawsuit challenging New York's sole use of the SAT to award scholarships. Federal Judge John Walker decided in *Sharif v. New York State Education Department* that this use of the SAT discriminated against women. One of the litigators of the case, Kary Moss of the Women's Rights Project of the American Civil Liberties Union, authored the article, which appeared recently in the *Berkeley Women's Law Journal* (1989-90, Volume 4, #2).

For a free copy of "Standardized Tests as a Tool of Exclusion: Improper Use of the SAT in New York," write to FairTest. The article is also a part of FairTest's "New York Scholarship Lawsuit Packet" which includes all of the documents from the case. The packet is available for $10.00 from FairTest. ❒

K-12 Testing

Genuine Accountability Campaign Grows

The **Campaign for Genuine Accountability in Education**, initiated by FairTest, continues to gain supporters. The National School Boards Association, the National Coalition of Advocates for Students, and Bank Street College of Education have added their voices to the growing national effort to largely replace standardized multiple-choice testing with new, performance-based means of assessment. Many individuals, particularly teachers, also continue to sign on. ❐

Speaking Engagements

FairTest staff members can be heard speaking at the following events in the next several months:

July 20-22: Education Activists Conference, Oberlin College, Ohio.

July 23: North Carolina Grade One & Two Assessment Training, Meredith College, Raleigh, NC.

October 3-7: National Chapter I Parents Assn., New Orleans.

October 2-5, National Association of College Admissions Counselors, Louisville, KY.

October 10-12: High Scope, Ypsilanti, MI.

October 13: Genesee Valley Assessment Conference, Rochester, NY.

November 15-18: National Association for the Education of Young Children, Washington, D.C. ❐

Writing, Reading Found Inadequate

The National Assessment of Educational Progress (NAEP) has reported that the majority of U.S. students are not adequate writers or readers for their grade level. The NAEP study *Learning to Write in Our Nation's Schools*, including the results of tests and surveys and released in June, found that the majority of students sampled from grades 4, 8 and 12 did not often write in class, did not like to write, and were unable to write adequate informative, persuasive or narrative passages. More than one-third of all students in grades 8 and 12, for example, reported that they were not required to write even two paragraphs per week.

A companion report, *Learning to Read in Our Nation's Schools*, reported that students in these grades did not read much in school or out, liked to read less as they grew older, and averaged only a low level of performance on the predominantly multiple-choice reading tests.

African American and Latino students showed substantially lower proficiency than whites in both reading and writing. In writing, for example, eighth-grade minority students scored, on average, lower than fourth grade white students; and twelfth grade minority students fared worse than eighth grade whites and only slightly better than fourth grade whites. Girls at all grade levels performed better in writing than boys.

Earlier studies had suggested that extending the writing sample time from 10-15 minutes to 20-30 minutes could benefit minority students. This NAEP study found that, to the contrary, longer periods were more likely to benefit white students. With longer time periods, whites improved significantly on four of the five tasks, African Americans and Latinos on only one of the five. The report concluded that those with greater writing skills were able to take advantage of the extended time, those with weaker writing skills could not.

This information, coupled with an earlier NAEP report showing little overall change in writing and reading ability from 1974 to 1988, has led critics to charge that the multiple-choice dominated curriculum of many schools has prevented students from improving their writing and reading ability. To encourage more writing and reading, a growing number of reformers are calling for changes in assessment to a performance-based system. Even the authors of the NAEP reports have criticized the predominant forms of instruction in reading and writing.

Additionally, many studies have shown that students in lower school tracks or ability groups receive a different curriculum from students in higher tracks. Lower track students do more drill and test preparation, and are less apt to write or read authentic and extended material. The higher percentages of African American and Latino students in lower tracks may partially explain the continuing performance gap between white and minority students.

Learning to Read in Our Nation's Schools and *Learning to Write in Our Nation's Schools* are available from NAEP, Box 6710, Princeton, NJ 08541; $14 each. ❐

K-12 Testing

Reformers Seek Halt to NAEP Expansion

A national coalition of education and civil rights groups, organized by FairTest, has issued a public letter calling on Congress, the Bush Administration and state governors to block the proposed expansion of the National Assessment of Educational Progress (NAEP) at least until current legal requirements for evaluating its effects are met. The letter, signed by more than 75 organizations and prominent individuals, also called on the educational leaders to direct the National Assessment Governing Board (NAGB) to target greater resources to developing alternatives to multiple-choice tests.

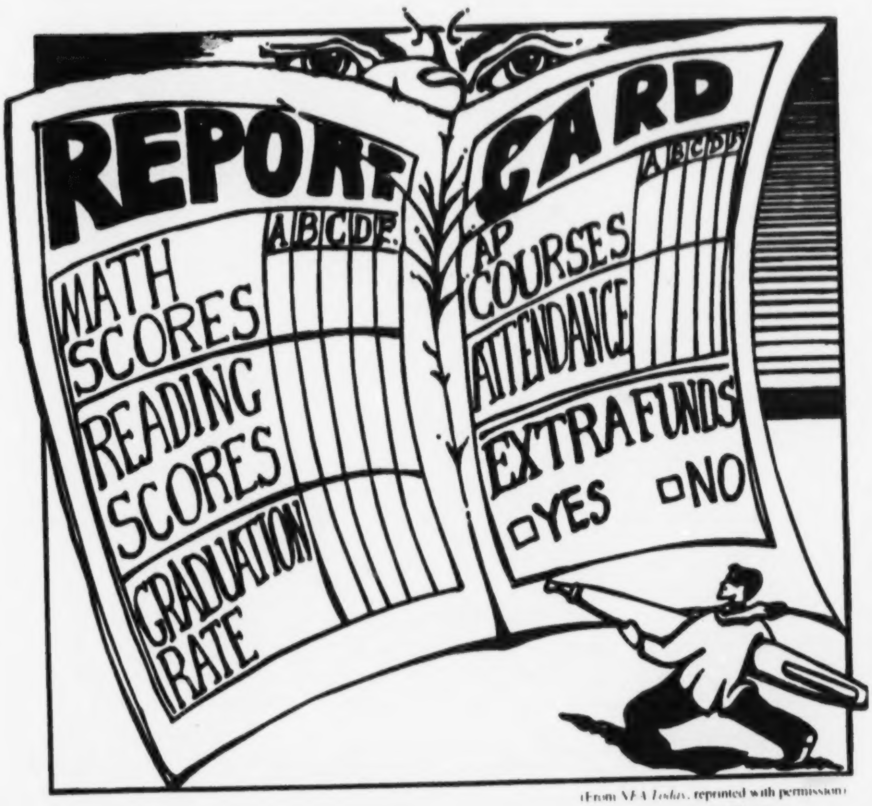
NAGB, which oversees NAEP, has come under fire in recent months for its proposals to expand NAEP and set national achievement levels on its tests, as well as its failure to proceed strongly toward adoption of performance-based assessment (see *Examiner*, Spring 1990). The signers of the letter informed educational policymakers that NAGB's recently adopted proposals to expand NAEP testing and set "achievement levels" for its exams were "dangerous" because they were adopted "without adequate consideration of the effects of the expansion or the proper role of assessment in education reform." The letter explained that expansion of NAEP must follow the final establishment of national goals and not be "allowed to dictate a national curriculum." The letter concluded, "Congress, the Administration and the Governors must weigh carefully the potentially harmful effects of more extensive testing and comparisons and ascertain that the dangers do not outweigh any possible benefits."

Among the concerns raised in the letter:

- NAGB adopted its proposal to expand NAEP before Congressionally-mandated studies on the validity and effects of pilot state-by-state comparisons were completed;
- NAGB's achievement level setting process, when combined with comparisons, may create a *de facto* national curriculum;
- By setting achievement goals based on multiple-choice tests, NAGB may define national educational goals in terms dictated by those narrow instruments;
- NAGB is moving too slowly in revising NAEP to reduce its reliance on multiple-choice items and move toward performance-based assessments;
- It is reckless to lift the ban on district-by-district or school-by-school test score comparisons without considering the educational consequences; and
- NAEP expansion is absorbing an ever-larger share of federal research and information funding without evidence that the results are worth the money.

Signers of the letter, in addition to FairTest, included the American

(continued on page 12)



(From NEA Today, reprinted with permission)

K-12 Testing

New York Campaign Aims To

A growing coalition of organizations in New York State is working together to ban the routine, mass use of standardized, multiple-choice type tests on young children from preschool through grade 2 (P-2). Organized by FairTest, the New York Public Interest Research Group (NYPIRG) and the American Reading Council (ARC), the New York Campaign for Genuine Accountability in Education has urged the State Board of Regents to issue a testing ban and is mounting a major public education effort to have districts stop P-2 testing. The Campaign also is calling on the New York State Education Department (SED) to make developmentally appropriate assessment methods available to schools across the state.

Although New York has few testing requirements for young children, a 1989 NYPIRG-FairTest study, *Failing Our Children*, documented that 75% of the districts surveyed test in kindergarten, 88% in first grade, and 93% in second grade. Twenty-four percent test prior to kindergarten. The districts use "readiness," "screening," "IQ," "aptitude," and "achievement" tests.

The Campaign Starts

Faced with the massive use of tests despite few requirements to administer them, and recognizing that testing young children is particularly harmful and inappropriate, FairTest, NYPIRG and ARC concluded that statewide action was necessary to correct the problem. In the fall of 1989, the groups sponsored a series of meetings, involving New York City area education and civil rights organizations, to develop a plan of action. Among the participants were the Academy for Educational Development, Advocates for Children, the N.Y. Urban Coalition, the Parents Coalition, the Public Education Association, the United Parents Association, and many individuals. The participants drafted a statement calling for the ban and developed a strategy to attain it.

As the New York City meetings progressed over the winter, NYPIRG and ARC began talking with members of the Board of Regents and staff from the SED. Many state agency officials expressed support for the goals of the Campaign.

A mailing list of hundreds of groups and individuals was compiled and used to inform and gather support. In reaching out across the state, the Campaign discovered several districts already seeking to minimize the amount of testing and its harmful impact. People in these districts - teachers, administrators and school board members - were eager to share their experiences and join in a wider effort.

Among the first major organizations to join the Campaign were the New York State Association for the Education of Young Children and the Genesee Valley Developmental Learning Group. These organizations work with young children from preschool through grade 2. Both have been seeking to implement developmentally appropriate curriculum, instruction and assessment, and both have extensive membership.

The Campaign's goals have also gained support from the state's largest teacher organizations, the New York City United Federation of Teachers (UFT) and the New York State United Teachers (NYSUT), both affiliates of the American Federation of Teachers (AFT). The UFT has approved a resolution stating that "a timetable should be set for phasing-out standardized, multiple-choice tests and replacing them with reliable, valid, and educationally appropriate alternatives."

(Illustration by Chris Demarest, reprinted with permission)

K-12 Testing

End Testing in Grades P-2

NYSUT set up a task force on assessment chaired by Antonia Cortese, its First Vice President. In a January statement to the Regents, Cortese explained that one of the three essential elements of school reform is to "[m]easure student success ... by assessing actual student performance rather than over reliance on multiple choice/short answer tests." The NYSUT task force has met twice to discuss current tests and performance-based assessment and will meet again to discuss its recommendations.

The state itself is considering the development of performance-based assessment. Education Commissioner Sobol has released a statement that includes a call for changes in assessment. And the state has used a hands-on grade 4 science test for the past several years.

To carry out Campaign work, FairTest and NYPIRG have received financial support from the Diamond Foundation and the New York Community Trust. But the effort's success also depends on substantial volunteer commitment.

Statewide Effort

Campaign organizers are reaching out to a broad range of groups around the state and have found wide support for the ban on P-2 testing (see list, p. 10). The Campaign is mobilizing to show this support to the Regents and other education policymakers. While developing a statewide network, the Campaign can simultaneously develop the strength in some school districts to win local bans. This process requires bringing many groups into the Campaign and having members of those groups become speakers and organizers. To this end, Joan Harris of NYPIRG and Susan Harman are making presentations at regional meetings.

At these sessions, volunteers from teacher, parent and community groups become part of the organizing effort. In a training workshop, they review essential information about why a ban is needed, familiarize themselves with Campaign material, and discuss how to carry the message to more people, get additional organizations to sign the Statement, and work toward a ban at the state and local levels.

In its outreach, the Campaign is aware of the particular needs and concerns of minority populations. Thus, the Campaign held meetings with organizations that represent

(continued on page 10)

New Assessments Needed for LEP Students

Increasingly, in-depth observation and documentation of limited-English proficiency (LEP) students is replacing conventional standardized tests for assessing their needs and achievements. English-language tests are unsuitable and invalid for use on these students and few standardized tests even exist in the native languages of many new immigrant children.

Rather than use test scores, a project in the Lea Elementary School in Philadelphia evaluated 50 LEP students using careful observation by trained staff over a period of weeks or months. Under the test-based decision process used in the past, most of the 50 would have been placed in special education. Under the observation system, only four were.

Alba Ortiz of the University of Texas estimates that 40 to 50 percent of LEP students in the U.S. have been mislabeled. "My sense is that an overwhelming number of LEP students who are called learning-disabled and put in special-education programs are normal, but underachieving, students for whom the system does not have an answer," she explained (*Education Week* April 18, 1990). Mislabeling leads to incorrect and often damaging placements in special education programs for learning disabled or emotionally disturbed students.

The Philadelphia observation process, developed in response to a lawsuit by Asian American families, is one of a number of new programs designed to avoid misplacement. California, New York City, Chicago, the District of Columbia and the Council for Exceptional Children are among those developing new assessment methods for LEP children. The observation-based process corresponds with recommendations of testing reformers to use documentation and performance-based assessments rather than current standardized exams. ❒

K-12 Testing

New York Campaign

(from page 9)

bilingual communities and with African American organizations. The Puerto Rican Legal Defense and Educational Fund co-sponsored the bilingual meeting, and U.S. Representative Major Owens co-sponsored the meeting with African American groups.

As campaigns at the regional and district levels gather momentum, activists will be going to their local school boards seeking a P-2 ban and the introduction of appropriate assessment. The local campaigns will also be pressing at the state level for the ban and for the dissemination of appropriate assessments and the staff development needed to educate teachers in their use. The outcomes of the Campaign should be a citizenry far more informed about testing, local assessment changes and finally a statewide ban. The effects of the Campaign should also influence testing reform at the higher grades.

The Campaign has developed educational tools to use in the effort. Currently, organizers have fact sheets on testing and authentic assessment developed by FairTest, an annotated bibliography, the Campaign Statement and the NYPIRG report. All are available from the Campaign, c/o Joan Harris, NYPIRG, 9 Murray St., NY, NY 10007; (212) 349-6460. ❐

Endorsers of New York Campaign

Advocates for African American Children
Advocates for Children
Alliance for Parental Involvement in Education, Inc.
The Arise Research Bureau
BCC Child Development Center
Beauty and the Beast Storytellers
Black Task Force on Child Abuse and Neglect
The Bronx New School Parent Teacher Association
Brooklyn New School
Buffalo State College Dept. of Elementary Education and Reading, Graduate Reading Faculty
Capital District Association for the Education of Young Children
Center for Collaborative Education
Champlain Valley Association for the Education of Young Children
Child Care Inc.
Diapering Alternatives
Early Childhood Education Council of New York City
Early Childhood Education Council of Western New York
Genesee Valley Developmental Learning Group
NAACP Legal Defense and Educational Fund, Inc.
New Program, PS 261
New York INTERFACE Development Project, Inc.
New York State Association for the Education of Young Children
New York State Child Care Coordinating Council
New York Urban Coalition
North Country Association for the Education of Young Children
Parent Association of PS 19Q
Public Education Association
Primrose School of Lincolndale, NY
Ravena Coeymans Selkirk Schools, Selkirk, NY
Richard C. Owen Publishers
River East School, New York City
Syracuse Association for the Education of Young Children
Westchester Mid-Hudson Pre-Kindergarten Directors
United Parents Association of New York City

Many individual educational leaders have also endorsed the Statement of the New York Campaign. ❐

(Illustration by Chris Demarest, reprinted with permission)

Commission Report

(from Page 1)

clear, multiple-choice tests are highly susceptible to error, bias and misinterpretation. Too many people have been wrongly denied major educational and career opportunities because of one faulty test score. FairTest encourages all those who support the Commission's recommendations to join with us to stop the damage stemming from over-reliance on standardized testing and to move towards a system of authentic assessment which promotes genuine accountability and equal opportunity."

Copies of *From Gatekeeper to Gateway* are available for $6 from the National Commission on Testing and Public Policy, McGuinn 529, Boston College, Chestnut Hill, MA 02167. ❐

(Illustration by Chris Demarest, reprinted with permission)

Excerpt from the Commission report:

The Case of Antonia Gonzalez

"Antonia, an A student in high school, dropped out in her senior year to get married. Years later, at age 40, she completed her high school equivalency requirements, enrolled in a junior college, and finished two years with a 3.4 grade-point average. She subsequently transferred to a university to pursue a degree in education, earning B grades in her first two upper-level education courses. Antonia then encountered a roadblock: she failed the "rising juniors" test, recently required by her state for admission to an approved program of undergraduate teacher education.

"Antonia became a statistic, part of the 66 percent Hispanic failure rate on the "rising juniors" test. Although she passed the mathematics section of the test on her first attempt, she failed the reading and writing sections. A year later, after taking every test preparation course she could, she passed the writing test but failed the reading test by a single point. In her third attempt she again failed the reading test by one point and was prohibited from taking further upper-level education courses until she passed.

"Antonia's initial opportunity to prepare for a career in teaching was denied on the basis of a single point on a single test despite evidence of her previous success in upper-level education courses. . . .She went on to tutor in a home for disturbed boys. Experienced educators judged her teaching there to be outstanding.

"Why was Antonia barred by a test from a profession for which she showed promise and which desperately needs teachers from her ethnic and linguistic background? Many approach this question by asking another: "Why did she fail the mandated test?" There are several possible answers to this question. Antonia was an older student. English was not her first language. Measurement error and culturally and linguistically biased test questions may have negatively influenced her scores. All these factors may explain why Antonia "failed" to hurdle the wall of her state's mandated test.

". . . Antonia did not fail the test so much as the testing policy of her state failed both her and the educational needs of her community. The policy ignored the fact that since any test is a fallible instrument, some scores will always turn out to be lower than they should. More importantly, when a test score alone is used as a bar to educational or employment opportunity, some capable people—especially people of ethnic and language minority backgrounds—are inevitably misclassified. When clear and direct evidence of a person's competence is available—in Antonia's case, both satisfactory performance in college and success on the job—a test score should not be allowed to present an absolute bar to opportunity. The way the test was used was thus not only illogical; it was unfair. It misdirected Antonia's education by wasting her time and money in test preparation courses. Further, it undermined her state's efforts to encourage minority students to enter the field of teaching.

"Three years of study have convinced this Commission that stories such as Antonia's are common to many educational and employment testing programs, especially those that use test scores alone to determine initial opportunities." ❐

K-12 Testing

Performance-Based Assessment

multiple-choice testing by schools will not be prohibited, the focus on performance-based assessments should discourage the use of multiple-choice instruments.

The new Kentucky assessments will complement statewide curriculum reforms and a new system of school-based decision-making. The state plans substantial expenditures for staff development to help teachers and administrators implement the changes. This year, Kentucky chose not to administer the Comprehensive Test of Basic Skills (CTBS), adding the funds originally earmarked for that test to those already available for developing new assessments. Foster plans to assemble a committee of experts in performance-based assessment to oversee the development and implementation of the new evaluation procedures. They would also monitor an interim system in which the state will develop and use "a testing program similar to the one used by the National Assessment of Educational Progress."

This year, Kentucky chose not to administer the CTBS, adding the funds originally earmarked for that test to those already available for developing new assessments.

(For more details on Kentucky's law and plans, including a copy of "The Kentucky Education Reform Act of 1990," write to Jack Foster, Secretary of Education, the Cabinet, State Capitol, Frankfort, KY 40601.)

Vermont

Vermont has approved funding for statewide portfolio assessments and for what the Education Department terms a "uniform" test. Portfolios are compilations of student work analyzed by teacher and student to demonstrate academic progress. Next year, portfolios in writing and math will be piloted on 10% of the state's fourth and eighth-graders. They will be adopted statewide the following year. Writing portfolios will be evaluated according to range of content, the depth of revision and the student's willingness to take a risk. The student's "best piece" will be chosen and also evaluated. In the future, portfolios in science, history and citizenship will be developed.

The "uniform" tests in math and writing, to be developed by Vermont teachers, will be administered to all students in grades 4 and 8. While some math items will be multiple-choice, most will probably be open-ended. "We don't want to repeat other people's mistakes," explained Ross Brewer of the Department of Education. One reason for including multiple-choice items is to enable comparisons with nationwide data, which still is based on multiple-choice standardized tests. In writing, each student will respond to the same prompts (writing assignments or questions). The prompts will be piloted next spring. Both writing and math tests will be implemented statewide in 1991-92, the same year as the portfolio assessments.

The state's accountability system will involve all three assessment techniques—uniform test, portfolio,

NAEP Expansion

(from page 7)

Association of School Administrators, the National Parent Teachers Association, the NAACP, the National Education Association, the National Association of Secondary School Principals, the Association for Curriculum and Supervision Development, The National Council of Teachers of English, the American Reading Council, the National Coalition of Advocates for Children, the National Council for the Social Studies, and the International Reading Association.

According to FairTest Associate Director Monty Neill, "FairTest is concerned that an expansion of NAEP could undermine the school reform movement and impose unwarranted federal control over education." Judith Thelan, President-elect of the International Reading Association observed, "Assessment needs to be structured to benefit students and teachers. The International Reading Association is deeply concerned about the misuse of large-scale test results which can extinguish the spirit of true classroom reform." The letter noted that the signers are "not writing to oppose the national assessment, but to help ensure that it plays a constructive, not harmful, role in reforming our nation's educational systems."

(For a copy of the open letter and a complete list of signers, write to NAGB Letter, FairTest.) ❐

K-12 Testing Spreads Nationally

(from page 16)

and best piece. Teams of trained Vermont teachers will use published criteria to evaluate random samples of portfolios in each school and teachers' assessments of student portfolios. School, district and statewide results of tests and portfolios will be made public.

The Department of Education is developing an RFP for a company to conduct the assessments. They are also looking for an assessment director. (Anyone interested should contact Ross Brewer, Vt. Dept. of Education, Montpelier, VT 05602. The publication, *Vermont's Assessment System: How Will It Work*, which contains a good discussion of writing portfolios, is also available.)

Arizona

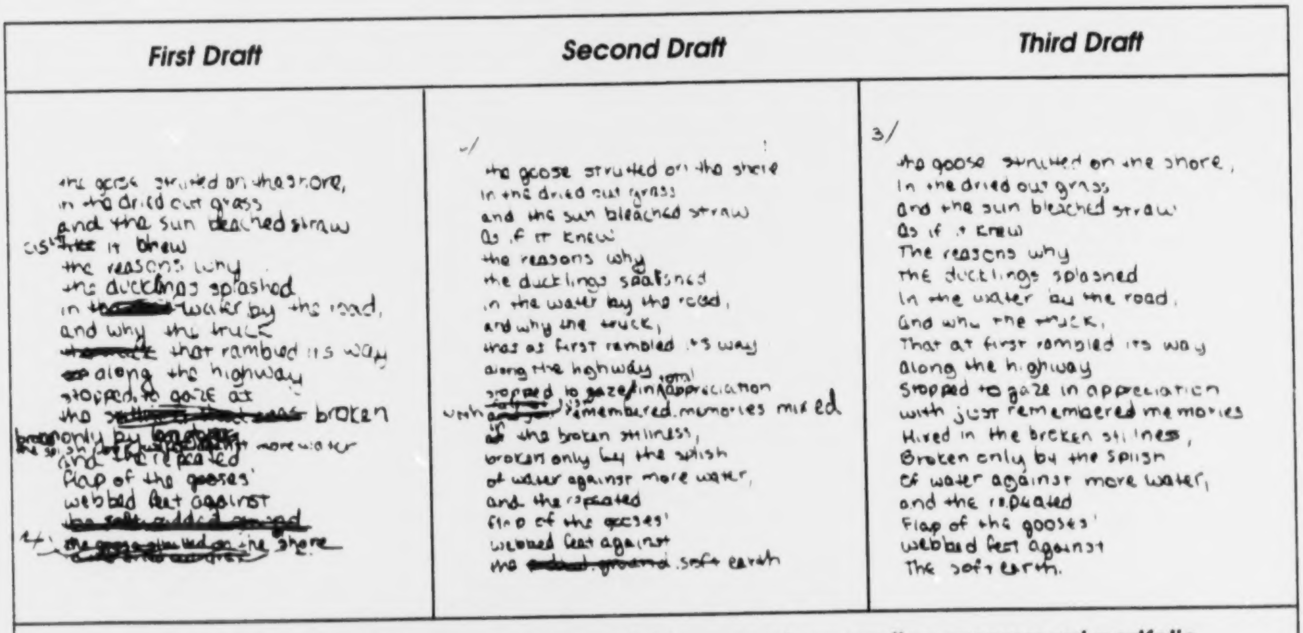
Arizona chose yet another route when an analysis of the Iowa Test of Basic Skills, the current statewide test, showed that it covered only about one quarter of the state's curriculum. Another analysis showed a tendency for teachers to teach primarily what was on the test. In response, Arizona's Department of Education proposed to develop performance-based assessments to evaluate student achievement in the rest of the curriculum, thereby also providing teachers with an impetus to teach the entire curriculum. The proposal became law this May.

For its new performance-based tests in reading, writing and math for grades three, eight and twelve, Arizona will develop three forms—A, B and C—at each grade level in each subject area. "Form A" is the pilot test and was administered to a sample of students this year. Next fall, a revised Form A will be given to each district in copyable form for teachers to use for instructional and assessment purposes.

For each grade and subject area, a number of assessments exist, one for each major objective. In the third grade, for example, there will be five reading, five writing (parallel to the reading) and eight math assessments.

Next spring, the state will make "Form B" available to the districts. Form B covers the same content areas in the same format as A, but comes in "easier" and "easiest" versions, suitable for use in the grades leading up to grades three, eight or twelve. This will enable schools to re-shape their curricula in all grades in line with the state's entire curriculum objectives.

(continued on page 14)



First Draft

the goose strutted on the shore,
in the dried out grass
and the sun bleached straw
as if ~~the~~ it knew
the reasons why
the ducklings splashed
in ~~[illegible]~~ water by the road,
and why the truck
~~the truck~~ that rambled its way
~~as~~ along the highway
stopped to gaze at
the ~~[illegible]~~ broken
~~broken~~ only by ~~[illegible]~~
~~the splash [illegible]~~ more water
and the repeated
flap of the gooses'
webbed feet against
~~the soft cold ground~~
4/ ~~the goose strutted on the~~ shore

Second Draft

2/
the goose strutted on the shore
in the dried out grass
and the sun bleached straw
as if it knew
the reasons why
the ducklings splashed
in the water by the road,
and why the truck,
that at first rambled its way
along the highway
stopped to gaze in appreciation
with ~~[illegible]~~ just remembered memories mixed
in ~~of~~ the broken stillness,
broken only by the splish
of water against more water,
and the repeated
flap of the gooses'
webbed feet against
the ~~cold ground~~ soft earth

Third Draft

3/
the goose strutted on the shore,
In the dried out grass
and the sun bleached straw
As if it knew
The reasons why
the ducklings splashed
In the water by the road,
And why the truck,
That at first rambled its way
Along the highway
Stopped to gaze in appreciation
with just remembered memories
Mixed in the broken stillness,
Broken only by the splish
Of water against more water,
and the repeated
Flap of the gooses'
webbed feet against
The soft earth.

Three drafts of a student poem, as they might appear in a reading assessment portfolio.

From "Portfolio Assessment: Sampling Student Work," by Dennie Palmer Wolf. Educational Leadership, *Vol. 46, No 7, April, 1989. Reprinted with permission of the Association for Supervision and Curriculum Development © 1989 by ASCD. All rights reserved.*

BEST COPY AVAILABLE

K-12 Testing

Performance-Based Assessment

(from page 13)

In the spring of 1992, the state will administer "Form C," an alternate form to Form A, as a test. Form C will be very similar to Form A. For example, a math problem on A asks students to design a playground; a similar problem on C might ask students to design an amusement park. Similar types of computation, problem solving and written explanation and description tasks will be involved in each assessment.

Every student in a given class will do the same assessment in a subject area. For example, one grade three reading assessment is to read and demonstrate comprehension of a personal experience narrative; each child in a class will do that assignment. Because the various assessments in each content area are not equatable, results will be reported separately for each assessment.

In June 1991, each district will be required to file an assessment plan to let the State Department of Education know how the district will observe, measure and be accountable for the competencies measured in grades three, eight and twelve in the years prior to those grades. The Department hopes that districts will use Forms A and B for instruction and internal assessment and that they will begin to implement portfolio evaluations at the school and district levels. "We don't always need to test to find out whether students are making progress in the competencies," observed Lois Easton of the Department.

Scoring of the performance-based assessments will be done by trained teachers. For the statewide testing, teachers who were trained during the pilot will train additional teacher-scorers. The pilot scoring demonstrated high inter-rater reliability.

While the state developed the specifications and prototypes for the new tests, the actual forms are developed by Riverside Publishing, makers of the Iowa Test of Basic Skills, which owns the copyright.

The Department hopes for major changes in curriculum and instruction through use of performance-based testing. However, the state will continue to administer a norm-referenced multiple-choice test each year in grades two through twelve. That test will now be given in the fall rather than the spring, reducing the pressure to teach to the test. Additionally, rather than administer an entire battery, each student will take only the comprehension part of the reading test, one other reading sub-test, and one language and one math sub-test. The result will be a one-half to two-thirds reduction in testing time.

(For more information, contact Lois Easton, Director of Curriculum and Instruction, Arizona Department of Education, 1535 W. Jefferson St., Phoenix, AZ 85007.)

Unresolved Issues

The standing-room-only audiences for performance-based assessment workshops at the ECS conference indicate the rapidly-growing interest in this method of assessment. But conference discussions also identified a variety of related issues which need further consideration. First, questions of potential race, class, culture and gender biases in the new assessments have only begun to be addressed.

Second, the relationship between classroom-based assessments, such as portfolios, and externally administered tests such as the California Assessment Program, which is becoming increasingly performance-based, has not been resolved. Vermont, for example, is focusing primarily on portfolios, while other states are concentrating on externally-administered exams.

Third, simply labeling a test "performance-based" does not make it a good test. In the laudable rush to change modes of assessment, much work will have to be done to assure quality.

Fourth, states and other agencies must do further work on technical problems to ensure that performance-based exams validly and reliably cover content areas; serve as tests worth teaching to; are not corrupted in the way that teaching to multiple-choice tests has corrupted both test results and instruction; and provide aggregatable data for state and national information.

Finally, new assessments cannot be meaningfully implemented without changes in curriculum, instruction and school structures. Such massive reforms will require changes in teacher and administrator training both in college and on the job. States still have a great deal of work to do in designing and implementing adequate methods of teacher education. Several states moving toward performance-based assessments have recognized the need for staff development. But much more needs to be done to ensure that this positive trend lives up to its promise. ❐

FairTest Publications

____ ***FairTest Examiner***, published quarterly. . . $15.00/year (individuals), $25.00 (institutions)

____ Back Issues of ***FairTest Examiner***, $4.00 each: Vol. __No. __(All for $32.00)

____ ***Standing Up to The SAT***, by John Weiss, Barbara Beckwith and Bob Schaeffer (192 pp.) . . . $6.95

____ ***Fallout From the Testing Explosion: How 100 Million Standardized Exams Undermine Equity and Excellence in America's Public Schools***, by Noe Medina and Monty Neill (3rd ed., 77pp.) . . $8.95

____ ***Sex Bias in College Admissions Tests: Why Women Lose Out***, by Phyllis Rosser and FairTest staff, 3rd ed. (60 pp.) $7.95

____ ***The Reign of ETS***, by Allan Nairn and Ralph Nader (550 pp.). . . $30

____ ***Beyond Standardized Tests: Admissions Alternatives That Work***, by Amy Allina with FairTest staff (18 pp.) $5.50

____ ***None of the Above, Behind the Myth of Scholastic Aptitude***, by David Owen (327 pp.) . . . $5.95

Contributions to FairTest

I agree there's no better way to be part of the testing reform movement than supporting FairTest! Enclosed is my tax-deductible contribution!*

____ **FairTest Associate** ($25.00--includes a free one-year subscription to the *FairTest Examiner*)

____ **FairTest Sponsor** ($50.00--includes *Examiner* subscription and *Standing Up to the SAT*)

____ **FairTest Sustainer** ($100.00 or more--includes above plus a 20% discount on publications)

____ **Other**

Publications Total $ ____

Contributions Total $ ____

Check Total $ ____

Name ____

Address ____

City ____ State ____ ZIP ____

238

Please make checks payable to FairTest and mail to: FairTest, 342 Broadway, Cambridge MA 02139

**Contributions to FairTest are fully tax deductible, except for the value of publications received.*

Performance-Based Assessment Spreads Across Nation

State plans to develop and use performance-based assessments are expanding rapidly. Unlike multiple-choice tests, performance-based assessments ask students to work on real tasks, such as writing or conducting experiments, thereby directly demonstrating knowledge and capabilities. At the June 1990 Education Commission of the States (ECS) Conference on Assessment, a number of states that have mandates to develop and use the new assessments agreed to form a consortium. The participants will share resources in developing and analyzing portfolios, open-ended test items and other forms of authentic evaluation.

This emerging consortium, to be coordinated by ECS, is only one of several being developed. Connecticut is working with six other states and the Coalition for Essential Schools to design performance-based assessments in math and science. Connecticut and other states are working with the Coalition, Project Zero at Harvard and the new Center for Diversified Learning and Assessment at Harvard, to secure financing for a parallel development in the arts and humanities.

The new ECS consortium could ultimately involve most of the states that are developing performance-based tests. Their combined efforts could become models for reforming federal testing programs such as Chapter I and the National Assessment of Educational Progress. What is emerging in the states, however, is not uniform: a variety of possibilities are being considered.

Kentucky

As part of a new law restructuring the state's public education system, Kentucky must "develop and implement a state-wide, primarily performance-based assessment program as early as the 1993-94 school year but no later than the 1995-96 school year." The assessments will be based on six basic educational objectives established by the legislation. Local districts will establish their own curricula using a statewide framework.

Jack Foster, Kentucky's Secretary of Education, told educators at an ECS workshop, "There should be no difference between what is taught and what is assessed. We shouldn't be able to tell them apart." Teachers and schools will be rewarded for improvement on the assessments, as well as on other measures, and face sanctions if there is no improvement. The assessments will focus on grades 3, 5, 8 and 12 and be used primarily to evaluate schools, not individuals. For example, the grade 12 exit test will not be used to deny diplomas.

In his talk, Foster acknowledged that the new assessments could play a role in grade-to-grade promotion decisions, but he expected retention to be minimal. While tracking is not barred, he thought the system's rewards and sanctions for having all children learn would lessen the creation of slow tracks. Though

(Continued on page 12)

FairTest®
National Center for Fair & Open Testing
342 Broadway
Cambridge, MA 02139-1802
(617) 864-4810
(617) 497-2224 (FAX)

Non-Profit
U.S. Postage
PAID
Cambridge, MA
Permit No. 53278

ADDRESS CORRECTION REQUESTED

Moving? Return your label with new address

FairTest
EXAMINER
The National Center for Fair & Open Testing Vol. 4, No. 4 Fall, 1990 ISSN # 0898-2511

FairTest Seeks Internal Revenue Service Investigation of National Merit Tax Exemption

In a letter delivered to the Internal Revenue Service in mid-October, FairTest has asked for a review of the National Merit Scholarship Corporation (NMSC) in light of NMSC's programs which discriminate against women and minorities.

In past years, FairTest has reported that nearly two-thirds of Merit Scholarships go to males, with only a small proportion awarded to blacks, Hispanics, Native Americans and recent Asian immigrants (*Examiner*, Summer 1989). Ample evidence demonstrates that NMSC's exclusive reliance on Premliminary Scholastic Aptitude Test/National Merit Scholarship Qualifying Test (PSAT/NMSQT) scores to determine eligibility for National Merit Scholarships unfairly reduces the chances for non-white and female students to win these prestigious and financially-valuable awards. Despite pressure from civil rights, feminist, and education reform groups, NMSC has refused to change its formula.

The FairTest letter to the IRS explained, "A basic tenet of American public policy is that educational institutions must not discriminate on the basis of race or sex. NMSC's exclusive reliance on scores from a biased exam... violates this fundamental policy of nondiscrimination. Therefore, NMSC should not be subsidized by federal money through tax exemption."

The letter noted that the "Internal Revenue Service has previously determined that institutions which fail to follow racially nondiscriminatory policies are not to be considered 'charitable'."

Only those students from each state who receive the highest PSAT/NMSQT scores are eligible to compete for Merit Scholarships. "Thus," FairTest argued, "rather than selecting from among the brightest and best college-bound students, NMSC picks finalists from a pool of outstanding *test-takers* by eliminating 99.5% of all applicants solely on the basis of their scores on a discriminatory test."

"Talented women and minorities,

(Continued on page 10)

Employment Testing Victories

Despite fears that the Supreme Court's *Wards Cove* ruling might block employment testing reform at all levels (*Examiner*, Summer 1989), advocates have won several major victories in the past few months.

Reversing precedents dating back to 1971, *Wards Cove* greatly increased the burden on plaintiffs seeking judicial remedies for discriminatory hiring practices. Nonetheless, U.S. government agencies have recently suspended a widely-used employment referral exam, instituted a much fairer process for filling Federal professional jobs, and soundly criticized so-called "integrity" tests used by many private employers.

GATB Suspended

In a decision with significant national impact, the U.S. Department of Labor (DOL) has discontinued its controversial General Aptitude Test

(Continued on page 12)

CONTENTS

Denise Carty-Bennia 2
Truth-in-Testing 3
NAEP & Summit 4
Biology Tests Blasted 5
Project Construct 6
Texas Parents Win 7
University Testing 8
Test Validity Lawsuit 11
Testing Reform Guide 16

TH 017 189

Denise Carty-Bennia 1947-1990: FairTest Board Chair

Photo courtesy of CUNY Law School

FairTest mourns the sudden death of Denise Carty-Bennia, our beloved friend, colleague, and Chair of our Board of Directors.

Denise helped found the National Center for Fair & Open Testing in 1985 at a conference of educators and civil rights activists held in Stony Point, New York. She served on the Board of Directors since the organization's incorporation and was its first chair.

A dynamic speaker and a tireless advocate, Denise was an inspiration to the cause of testing reform. At her last public appearance on behalf of FairTest, she testified movingly at a hearing in New York on the bar examination. Denise described the bar examination's discriminatory impact on minority students from her more than a dozen years of experience as a lawyer and law professor.

In addition to her positions as a tenured professor of law at Northeastern University Law School and visiting professor at the City University of New York Law School, Denise was involved in countless struggles for human rights. These ranged from co-authoring a brief submitted to the U.S. Supreme Court on behalf of 900 law professors in the abortion rights case, *Webster v. Reproductive Health Services*, to co-chairing the National Conference of Black Lawyers, to serving on the board of directors of the National Employment Law Project. She also was a Research Fellow of the Institute for the Study of Educational Policy at Howard University and received a Charles Revson Fellowship. Denise graduated from Barnard College and Columbia University School of Law.

Denise contributed enormous time, energy and savvy to FairTest, and for that we are all grateful. FairTest will always remember her friendship and love, which she gave selflessly. The entire testing reform community will miss Denise Carty-Bennia.

To memorialize her years here, FairTest is creating a program that celebrates the principles to which she dedicated her life. The **Denise Carty-Bennia Internship** will provide students with an opportunity to work on special projects devoted to eradicating racial, ethnic, and gender bias in testing. For more information on participating in this program, or contributing to its support, please write to FairTest. ❐

FairTest

National Center for
Fair & Open Testing
342 Broadway
Cambridge, MA 02139-1802
(617) 864-4810 (617) 497-2224 (FAX)
FairTest EXAMINER, Fall 1990

Staff:
Monty Neill, Associate Director; **Laura Pierce**, Administrative Assistant; **Bob Schaeffer**, Public Education Director; **Cinthia Schuman, Esq.**, Executive Director; **Jamie Souder**, Publications/ Office Manager; **Sarah Stockwell**, University Tests Director; **Veda Wright**, Field Organizer.

Board of Directors:
Dr. Pamela George, North Carolina Central University; **Dr. Susana Navarro**, California Achievement Council; **Paul Pottinger**, The Equipment Financing Group; **William Robinson, Esq.**, Dean, D.C. School of Law; **Bob Schaeffer**, Treasurer; **Steve Solomon, Esq.**, Proskauer, Rose, Goetz & Mendelsohn; **Chuck Stone**, *Philadelphia Daily News;* **Ron Vera, Esq.**, Best, Best & Krieger; **John Weiss**, Public Policy Initiatives, President.

National Policy Panel:
Dr. Leonard Beckum, Emory College; **Dr. Beverly Cole**, NAACP; **Dr. Harold Dent**, Association of Black Psychologists; **Arnold Fege**, National PTA; **Dr. Howard Gardner**, Harvard University; **Dr. Asa Hilliard**, Georgia State University; **Dr. Lucinda Lee Katz**, University of Chicago Laboratory Schools; **U.S. Rep. Matthew Martinez** (D-CA); **Deborah Meier**, Central Park East Secondary School; **Sara Melendez**, University of Bridgeport; **Aurelio Montemayor**, Intercultural Development Research Association; **Harry Pachon**, National Association of Latino Elected and Appointed Officials; **Alfredo de los Santos**, Maricopa Community Colleges; **Eleanor Smeal**, Fund for a Feminist Majority; **Theodore Shaw, Esq.**, NAACP Legal Defense Fund; **Dr. Gail Thomas**, Texas A & M University; **U.S. Rep. Ted Weiss** (D-NY).

Funding for FairTest's work comes from the Golden Rule Corporation, the Ford Foundation, the Rockefeller Family Fund, the Philip M. Stern Family Fund, the Aaron Diamond Foundation, the Edward Hazen Foundation, the Lilly Endowment, the New York Community Trust, and many individual contributors.

Truth-in-Testing Legal Brief Filed

In response to efforts by test producers to overturn New York's landmark Truth-in-Testing law, FairTest and ten civil rights, women's rights and educational leaders have filed an *amicus curiae* brief in the appeal of *Association of American Medical Colleges v. Cuomo*. That case struck down Truth-in-Testing--which requires testing companies to publicly disclose university admissions test questions and answers--on the ground that the New York statute violated federal copyright law (Examiner, Spring1990). While the decision applied only to the makers of the Medical College Admissions Test (MCAT), other testing companies have begun using the decision to rid themselves of Truth-in-Testing requirements.

Already, the Educational Testing Service, College Board, Graduate Record Examination Board, Graduate Management Admission Council and Test of English as a Foreign Language Policy Council Board have sought a preliminary injunction in federal court to stop New York from enforcing the Truth-in-Testing law against their exams. The companies have agreed to partially comply with the law, pending the outcome of *Association of American Medical Colleges v. Cuomo*.

In the *amicus* brief, FairTest argued that Truth-in-Testing was enacted "in response to concerns that standardized testing companies had become monopolies that controlled exams having great impact on the lives of many New Yorkers, but were withholding information needed to ensure their accountability to test-takers and the public." Evidence that standardized tests contained questions that were "ambiguous or unfair, questions having more than one correct answer, questions that could shock or offend test-takers, and questions that might lead to biased results" led New York to enact the disclosure law.

Truth-in-Testing has helped open the door to testing reform. One indicator of its importance is the range of groups and individuals that signed onto the FairTest brief. These include: the Center for Women Policy Studies, Deputy Assembly Speaker Arthur Eve (D-NY), Fund for a Feminist Majority, Golden Rule Insurance Company, Mexican American Legal Defense and Educational Fund, National Education Association, National Women's Law Center, New York Public Interest Re-search Group, NOW Legal Defense and Education Fund, and Public Citizen.

The case of *AAMC v Cuomo* will be argued in the Second Circuit Court of Appeals later this fall.

Thanks to Debevoise & Plimpton, Public Citizen

FairTest is grateful to the law firm of Debevoise & Plimpton, for providing excellent *pro bono* counsel in the Truth-in-Testing litigation, and to the Public Citizen Litigation Group for legal assistance. Special thanks to the legal team of Bruce Keller, Lorin Reisner, Alan Morrison, Geoffrey Kaiser and David Vladeck. ❒

NTE Scoring Errors Show Need for Truth-in-Testing

Scores on the National Teachers Examinations (NTE) for 49,000 test-takers have been revised upward after a review revealed that six items had more than one correct answer, according to *Education Week* (October 10). The review of the multiple-choice tests was conducted by the test-maker, Educational Testing Service (ETS).

Two of the items were on core-battery tests administered between October, 1984, and June, 1990. The score increases resulted in 998 examinees now scoring above the cut-off scores set by the states in which the test-takers had planned to teach. Another 166 examinees had scores raised above the cut-offs on the physical education specialty-area tests. Four items on this test, administered between November 1989 and February 1990, were determined to have more than one correct answer.

FairTest heard from one New York woman who was denied a teaching license five years ago because she failed by one point to attain the state's cut-off score on the NTE core battery. Several years later, she passed the exam and finally began to teach. Recently, she received notification from ETS revising her first scores upward. Who, she wanted to know, would compensate her for the years of lost income, emotional distress, loss of esteem, and professional damage caused by the inaccurate test?

If Truth-in-Testing were in place for the NTE, these flawed items could have been challenged by test-takers and detected much earlier, preventing or lessening the hardship inflicted on examinees. Further, this situation demonstrates the arbitrary and harmful effects of implementing cut scores on exams. ❒

K-12 TESTING

Battle Looms Over Post-Summit Panel

The panel appointed by the White House and the National Governors Association (NGA) to oversee implementation of the nation's educational goals is facing a strong challenge from Congress. The panel, set up after last fall's Charlottesville, VA "Education Summit," will not include any educators or other members of the public. This composition was labeled "ridiculous" by Bill Goodling (R-PA), ranking minority member on the House Education and Labor Committee. A Senate committee has approved legislation creating an independent monitoring body.

The 14-member administration/governors' panel will include six governors (three from each party) and four federal officials appointed by the President. Four Congressional leaders (two from each party) will serve as *ex officio*, nonvoting, members. The panel is supposed to work by consensus, with a 75% majority required to break any deadlock.

Congressional leaders have charged that such a panel of political figures cannot ensure accountability. The Senate proposal (S.3095) would create an 18-member "National Council on Educational Goals." Ten members would be appointed by Congress, two by the President, and six by the NGA. The Congressional appointees would have expertise in education or assessment and could not hold state or federal elected office. If it passes the Senate, the legislation would be included in House-Senate conference discussions on education before adjournment in October.

That leaves the prospect of two competing panels trying to monitor the nation's educational progress. While the Bush-NGA panel lacks any legal authority, a Congressionally-mandated panel would have the full force of law and federal funding. ❐

New Education Congress Supports Authentic Assessment

At its first meeting, August 26-29, in Aspen, Colorado, most participants in a "Curriculum Congress" agreed that schools should stop relying on standardized test scores as measures of academic quality. Organized by the Education Commission of the States and Collaboratives for Humanities and Arts Teaching, the newly formed organization is composed of 20 groups representing teachers of subjects such as science, mathematics, English and social studies.

One of the recommendations endorsed by most participants was that schools should replace standardized tests with authentic ways of assessing academic progress. ❐

NAEP Achievement Standards

Despite continuing criticism of its efforts, a panel of the National Assessment of Educational Progress (NAEP) has met to establish achievement standards for last winter's NAEP math assessment. The 71-member panel of educators, public officials, and business people is charged with setting standards for three levels of achievement ("basic," "proficient" and "advanced") for grades 4, 8 and 12.

The National Assessment tests in reading, math, and other subjects are periodically administered to a national sample of students. The program is supervised by the National Assessment Governing Board (NAGB).

The standard-setting plan has received criticism from many sources (see *Examiner*, Spring 1990). For example, critics charged that the achievement standards are being driven by predominantly multiple-choice tests constructed before achievement levels or goals were even defined. Thomas Romberg of the University of Wisconsin told *Education Week* (Sept. 5, 1990) that the standards "may just add to the problem of getting kids to work on more complex tasks" since "teachers may spend more time on drill and practice on trivial stuff" in an effort to raise scores on the test.

Some panelists found it difficult to distinguish between achievement levels based on test items, particularly because of the lack of advanced-level items. Other critics have charged that the achievement information will lack educational utility, either because the information will not be instructionally useful to teachers or because it will be ignored. ❐

(For more NAEP news, see p. 14.)

K-12 TESTING

Biology Tests Blasted

Tests are "impediments to implementing curricular change," says the National Research Council (NRC) in *Fulfilling the Promise: Biology Education in the Nation's Schools.* The NRC was organized by the National Academy of Sciences to help advise the government and the public on scientific matters. Their just-released book describes in detail many problems caused by use of multiple-choice standardized tests and urges changes in biology assessment.

Among the reports findings and conclusions:

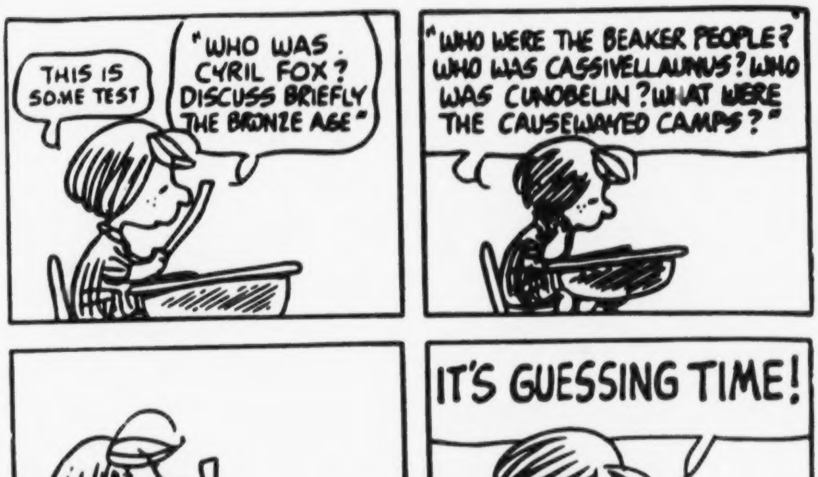
"Standardized tests use multiple choice items as the major or sole source of information.... Such tests are limited in determining achievement, in that they measure...only a small sample of factual information. Tests that rely on recall do not reflect the conceptual nature of biology, and they reduce and distort testing for concepts by testing for words and definitions instead. Standardized questions that merely require simple recall or name association are of little value to either teachers or students.... Standardized tests rarely assess the application of knowledge, problem-solving, or the ability to think through issues that involve biological understanding....

"With the recent press for accountability from national, state, and local agencies, testing programs are beginning to drive the curriculum.... Dependence on norm- or criterion-referenced multiple-choice tests as they are currently constructed, either for grading students or for analyzing and creating educational policy, tends to misdirect students' habits of study. Rather than mastering concepts, students believe that recognizing terms in a multiple-choice format is the appropriate educational goal.... In the long run, the impact of current modes of testing on enduring skills and strategies for learning will be inimical to reform....

"A major effort is needed to develop, test, and publish model examinations.... They should... [involve] students in thinking and reasoning.... States and school districts should move swiftly to recognize the need for new tests....

"The process of evaluation cannot be addressed in isolation from what is in textbooks, how teachers are taught, the conditions under which teachers work, and public consensus about the goals of science education.... [T]he purposes of norm-referenced testing do not serve the goal of improving teaching and learning. We need a national perspective concerned more with evaluating the effects of curricula, teaching methods, and materials than with ranking the performance of individual students.... Existing institutions with a role in testing are not designed to pursue that objective by themselves." (Selections from pp. 40-48.)

***Fulfilling the Promise* is available from National Academy Press, 2101 Constitution Ave., NW, Washington, D.C. 20418; prepaid, $14.95 plus $2.00 postage.** ☐


THIS IS SOME TEST
"WHO WAS CYRIL FOX? DISCUSS BRIEFLY THE BRONZE AGE"
"WHO WERE THE BEAKER PEOPLE? WHO WAS CASSIVELLAUNUS? WHO WAS CUNOBELIN? WHAT WERE THE CAUSEWAYED CAMPS?"
IT'S GUESSING TIME!
SCHULZ

BEST COPY AVAILABLE

Missouri's "Project Construct"

Alternatives to testing young children are being implemented in 50 pilot districts across Missouri as part of Project Construct, a voluntary program for Missouri schools. This fall, 200 teachers are participating in Construct's Formative Assessment Program, and 100 teachers and administrators are working with the University of Missouri's Center for Educational Assessment (CEA) on the Summative Assessment Inventory for Project Construct. This past summer, 260 teachers participated in a training institute conducted by the state's Department of Education.

In the Formative Assessment Program, direct observation of student activity is the primary basis for assessment. Teachers are encouraged to choose from a variety of methods to document a child's progress. Observation worksheets and record-keeping documents provide teachers with ways to record and collect data about student progress toward the twenty-eight curricular goals of the project. Portfolios of a child's work and student interviews can also be used. Periodically, teachers write up "summative assessments" to indicate the knowledge and capabilities attained by the child at that point in time.

The curriculum and assessment goals and materials included in Project Construct are based on a theory called Constructivism. Constructivism is based in research which shows that children learn by actively constructing knowledge as a result of interactions with their physical and social worlds.

Project Construct's approach to educating and assessing 3 to 8 year-olds corresponds to the developmentally appropriate needs of this age group. Standardized, multiple-choice tests, however, are rooted in a view that learning is based on passively accumulating facts.

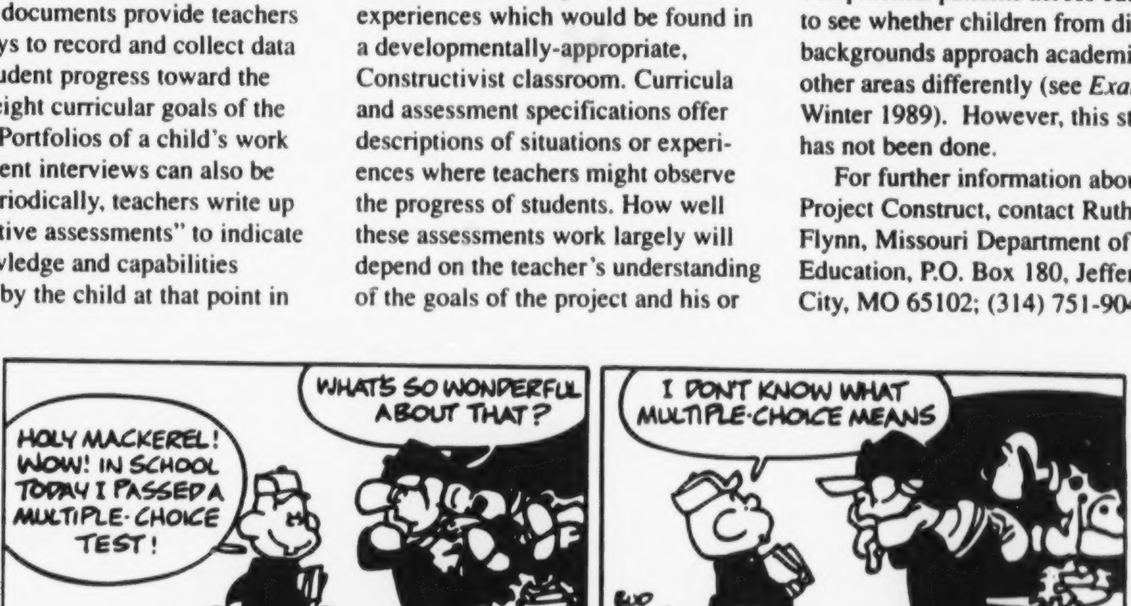
Project materials include detailed descriptions of the goals which come under four domains: sociomoral, cognitive, representational and physical development. Materials describe 16 learning and assessment experiences which would be found in a developmentally-appropriate, Constructivist classroom. Curricula and assessment specifications offer descriptions of situations or experiences where teachers might observe the progress of students. How well these assessments work largely will depend on the teacher's understanding of the goals of the project and his or her skills in observation, making the education of teachers central to the work of Project Construct.

Other than training teachers, no additional money was required to fund Project Construct. Development of curriculum and assessment came out of existing programs in the Department of Elementary and Secondary Education and the Center for Educational Assessment.

Parent involvement is an important aspect of Project Construct and will tie in with the state's Parents as Teachers objectives. A questionnaire is being developed that asks parents such questions as: "What do you think your child is good at doing?" The questionnaire will assist teachers in their work with the children who participate in Project Construct.

Samples of students on whom the project was piloted included a racial and ethnic mix of children. Initially, Project Construct planned to study developmental patterns across cultures to see whether children from different backgrounds approach academic or other areas differently (see *Examiner*, Winter 1989). However, this study has not been done.

For further information about Project Construct, contact Ruth Flynn, Missouri Department of Education, P.O. Box 180, Jefferson City, MO 65102; (314) 751-9046. ❒


HOLY MACKEREL! WOW! IN SCHOOL TODAY I PASSED A MULTIPLE-CHOICE TEST!
WHAT'S SO WONDERFUL ABOUT THAT?
I DON'T KNOW WHAT MULTIPLE-CHOICE MEANS
BUD BLAKE 2-17

K-12 TESTING

Texas Parents Stop Kindergarten Retention

Grassroots activism has persuaded the Texas Board of Education to bar all schools in the state from retaining children in pre-kindergarten and kindergarten. Responding to parents and teachers who wanted to see transitional classes eliminated, the Texas Education Agency advised school districts that transition year classes before first grade amount to retention, and are thus banned. The disproportionately large number of children of color placed in transitional classes helped spur the Board's decision.

Pressure on schools to raise standardized test scores has been one of the major factors in the rapid growth of transitional year classes nationally. Though transitional classes can sometimes be developmentally more appropriate than regular classes, research indicates that the overall harmful effects of this form of retention far outweigh the temporary benefits.

Readiness Testing Unreliable

Decisions to place children in transitional programs are often made on the basis of readiness tests, such as the Gesell, which have been criticized as extremely inaccurate and biased against students from low-income families and minority groups, as well as boys. Kids, Education and You (KEY), a Denver City, Texas, parent group that led the Texas effort to bar transitional programs, pointed out that the Board's ruling will have the effect of eliminating the primary reason for using readiness tests.

The National Association for the Education of Young Children (NAEYC) calls school readiness tests unreliable measures which should not be used to delay school entry or assign children to special classes. A position statement adopted this summer by the group calls for schools to be responsible for meeting the needs of all children as they enter. Chronological age should be the only determining factor for entrance to school, according to the NAEYC. ❐

NAESP Recommends Limiting Tests

Acting in agreement with a growing number of organizations and early childhood specialists, the National Association of Elementary School Principals (NAESP) has issued a guide which warns against using standardized tests to delay school entry, gauge progress or make placement or promotion decisions about young children. Such tests "do not reflect developmental theory and practice," concludes NAESP.

Instead of reliance on test scores, NAESP recommends use of a range of developmentally appropriate assessment techniques for 3 to 8 year-old children. NAESP also endorses learning centers, low child-adult ratios, use of alternative grouping strategies, and early childhood training for teachers. ❐

Delaware Reduces Testing

In a recent revision of state laws governing testing, Delaware cut its state-mandated annual testing by more than half. The state dropped testing for grades 1, 2, 4, 5 and 7, but continues to require a norm-referenced test in grades 3, 6, 8 and 11. Previously, the state required students in grades 1-8 and 11 to take the Comprehensive Test of Basic Skills (CTBS) each year. Special education students, even those who have been mainstreamed, are exempt from the mandated testing.

The changes were urged by activist groups representing teachers of young children, curriculum specialists in reading and language arts, and special education teachers. These groups campaigned at the state and local levels against any standardized tests.

Some Delaware districts, such as the Christina district covering parts of Wilmington and nearby suburbs, are now introducing performance-based assessments, particularly for young children. In Christina, the curriculum in the early grades is developmentally based; that is, founded on modern research on how children develop and learn. The assessment was constructed to meet teachers' needs for educationally useful information. Teachers are now using portfolios, samples of student work, and checklists. ❐

UNIVERSITY TESTING

New GRE Guidelines: Lip Service or Real Reform?

Due in part to "increased national attention to the use and misuse of test scores throughout the American educational system," the Graduate Record Examinations Board has announced that it will no longer help defend institutions which become involved in litigation because of "inappropriate score use." The new policy begins this fall.

The new standards are considerably stronger than policies for other tests, such as the SAT, LSAT or GMAT. Misuses which will no longer be defended include using GRE scores "in isolation" and using a GRE cutoff score "as a sole criterion for denial of admission."

In a cover letter accompanying the *1989-90 Guidelines for the Use of GRE Scores*, which was sent to test score users, GRE Board Chair Jasper D. Memory said, "Once the new guidelines become effective, neither the GRE Board nor Educational Testing Service will support any institution in any such litigation where it appears that the institution has engaged in inappropriate score use as specified in the Guidelines." Memory pointed out, "Institutions may miss good students if they rely only on standardized test scores."

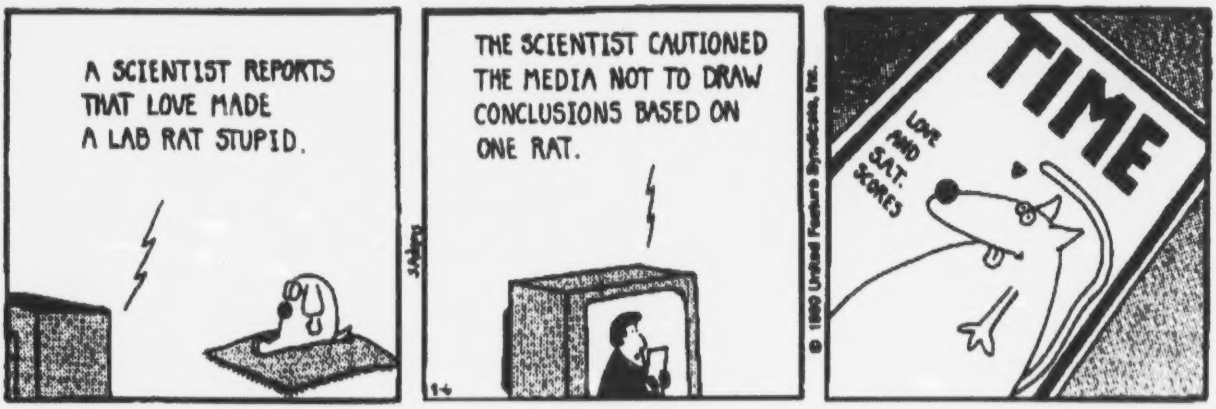
The actual impact of these new warnings on graduate schools, and whether the GRE Board and ETS will stand by their proclamations, remain to be seen. ❐

Test-Taker Wins Suit Against ETS

In the first known successful suit against the Educational Testing Service on issues concerning testing conditions, a New York City high school student has won the right to retake the SAT and have her earlier scores thrown out because of disruption at the test site. Emily Mindel, who had taken the SAT in March of this year, took the test again in May after having received coaching. Mindel stated there were a number of distractions at the May administration, including a proctor taking her test booklet away to show something to another test-taker, a small child repeatedly walking through the room during the test, and the proctor questioning her about marks in her test booklet during the exam.

Judge Beverly S. Cohen, of the New York Supreme Court, found the claims compelling because of the disruption and because Mindel had taken the May test to increase her scores in order to apply to several very selective schools under "early decision" programs. The Judge decided that Mindel should not "be forced to submit test scores attained on the May 5, 1990 test or wait and take the test in November, 1990 and thereby have her test results reported so late in the process that she may miss the deadline for the early decision process."

According to ETS' response to Mindel's charges, the disturbances were not in any way "beyond the normal range of events occurring at test sites whenever the tests were given." While this statement certainly raises alarming questions about the overall quality and fairness of ETS' testing conditions, test-takers currently have no recourse other than expensive legal action. A better alternative would be "due process" protections, which would provide for rapid resolution of disputes through a neutral party. But ETS successfully lobbied in the New York legislature this year against a bill to do just that. ❐


A SCIENTIST REPORTS THAT LOVE MADE A LAB RAT STUPID.
THE SCIENTIST CAUTIONED THE MEDIA NOT TO DRAW CONCLUSIONS BASED ON ONE RAT.
S.Adams
1-4
TIME
LOVE AND S.A.T. SCORES

© 1990 United Feature Syndicate, Inc.

UNIVERSITY TESTING

Lawsuit Charges ACT Cut Scores Promote Segregation

Though it was overturned on appeal, a recent federal court decision concerning the Mississippi higher education system has again raised controversy over the use of arbitrary minimum test scores to automatically reject applicants which critics charge can lead to racial segregation.

In *Ayers v. Allain*, handed down in February, a three judge panel of the Fifth Circuit U.S. Court of Appeals found that Mississippi's colleges and universities were illegally segregated and that test scores had played a role in the promoting of this segregation. According to the panel's decision, a cut-off score of 15 on the ACT was implemented as a requirement for admission to the historically white public universities in Mississippi by the state's board of higher education "soon after the court ordered admission of James Meredith to the University of Mississippi in 1962 because it deterred black enrollment." The cut score effectively ruled out most blacks, whose average score at the time was seven, while admitting most whites, who averaged 18. As a result, the University has remained overwhelmingly white: even today only 7.6% of the student body is African American.

This fall, the full Fifth Circuit Court overruled the panel, but plaintiffs are considering taking the case to the U.S. Supreme Court for final resolution. In the latest decision, the Court rejected the plaintiff's charge that black students were discriminated against by Mississippi's use of the ACT in admissions decisions.

Despite the recent decision, civil rights organizations maintain that the ACT continues to be used as a tool to reinforce segregation in Mississippi. Currently, historically white institutions require an ACT score of 15 for automatic admission, while historically black institutions use a cut-off score of 13. Although both types of schools allow exceptions to the cut-off, the policy is little known and rarely used at the historically white institutions, but frequently used at historically black institutions. Thus, the test continues to channel African Americans, who now average 4.3 points lower on the ACT than whites, into historically black institutions and whites into the better-funded, historically white schools.

Instead of waiting for another court decision to rectify Mississippi's misuse of the ACT, the American College Testing Service should stop this policy, which violates its own test use guidelines. Simply withholding test scores from the University of Mississippi system, or any other institution which applies cut-off scores, would stop such misuses. ❐

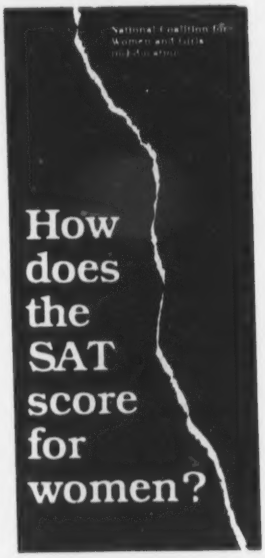
New Gender Bias Brochure Available

The testing task force of the National Coalition for Women and Girls in Education, co-chaired by FairTest and the Center for Women Policy Studies, has just published *How does the SAT Score for Women?* The new brochure addresses questions such as "Why do women and men score differently on the SAT?" "What is the impact of lower scores on women?" and "What can you do?" Copies of the brochure are available from FairTest for $1.50 (write for bulk prices). ❐

New Jersey Students Boycott Test

Widespread student and faculty protests have disrupted the first administrations of New Jersey's General Intellectual Skills (GIS) test, produced by the Educational Testing Service (ETS). Only three students at Glassboro State College and six at Kean College took the new volunteer "basic skills" test, rather than the 200-300 state officials had hoped for from each school. Concerns about bias, due process for test-takers, and the validity and reliability of the non-multiple choice test, as well as fears that the state will use it to rank public colleges and universities, sparked the protest.

Students also objected to the use of "incentives" to recruit test-takers at some schools. For example, officials at Rutgers University, where 315 students took the GIS, offered all test-takers a $20 bookstore voucher and awarded $250 to the student with the highest score. ❐



National Coalition for Women and Girls in Education

How does the SAT score for women?

National Merit

(Continued from page 1)

many at the top of their classes, are thereby eliminated from the competition by a single test score, which even the test-makers say should never be used in isolation," the letter continued. "As a result, nearly two-thirds of National Merit awards go to males year after year, as National Merit openly admits."

National Merit refuses to make the ethnic breakdown of scholarship winners public. However, because of the large score gap between whites and minorities on the qualifying test, it is unlikely that Blacks, Hispanics, Native Americans and members of recent Asian immigrant groups win more that a few percent of National Merit Scholarships.

Last year, Federal District Court Judge John M. Walker found the Scholastic Aptitude Test (SAT), from which the PSAT/NMSQT is derived, to be discriminatory when used to select scholarship winners in a New York State program similar to National Merit. Judge Walker said, "The probability, absent discriminatory causes, that women would consistently score 60 points less than men is nearly zero."

Testing reformers argue that better ways to award merit scholarships already exist. For example, the Citizens Scholarship Foundation in St. Peter, Minnesota, which administers many scholarship programs for communities, corporations and other organizations, considers factors such as high school record, honors courses taken, and leadership, rather than test scores. The states of Illinois and California also have large merit scholarship programs which do not rely on test scores to select winners.

As of this writing IRS officials have not yet responded to the FairTest letter. ❐

TEST
Campbell

"I knew the answers. I just couldn't retrieve them from my memory bank."

The Prices Of Test Secrecy

Allowing individuals other than test publishers to examine and screen tests can help improve test quality, reduce bias in the instruments, and ensure that they are worthy tools for assessing the health of our schools, according to a recent report of the Educational Technology Center at Harvard University's Graduate School of Education.

The report--which features a compilation of articles by educational scholars--examines the social, intellectual and psychological costs of using test instruments that are not subject to public inspection and discussion. It recommends that large, publicly-available data bases of reviewed test problems be accessible through libraries and bookstores.

"Assessment inevitably influences instruction," says the report. "If these data bases of problems were publicly available and school authorities published the criteria and procedures they used to select problems for accountability assessment, then the very criteria and procedures...would become one way of holding the school accountable and judging its quality. There would be continuing pressure from colleges and universities on the one hand from commerce and industry on the other to improve the problem collections. In such a climate the mediocrity of present-day multiple-choice items would not likely survive. Instead...we would see a growth in the number of problems that require students to think as well as to remember, to analyze as well as to regurgitate."

The Prices of Secrecy: The Social, Intellectual, and Psychological Costs of Current Assessment Practice, edited by Judah L. Schwartz and Katherine A. Viator, is available for $10.00 from the Educational Technology Center, Harvard Graduate School of Education, Cambridge, MA 02138. ❐

K-12 TESTING

Suit Challenges Validity Procedures

Parents who teach their children at home have filed a class action suit in South Carolina challenging a state validity study on the use of the Education Entrance Exam (EEE) to determine which parents can legally teach their children at home. The test is most often administered in South Carolina to sophomores seeking to enter teacher training programs.

When South Carolina passed a law requiring parents to pass the EEE to be allowed to teach at home, the exam's manufacturer, the Educational Testing Service (ETS), told the state it should conduct a validity study before it could properly use the test for this purpose. In response, the state conducted a traditional content validity study: 33 individuals were assembled to review the test's content. This method, sometimes referred to as BOGSAT (Bunch of Guys Sitting Around a Table), has been criticized by testing reformers as inadequate. The state-funded panel, however, deemed the test valid.

The suit challenging the validity study, *Lawrence, et al., v. South Carolina Department of Education, et al.*, was filed in state court by home-schooling parents, supported by the Home School Legal Defense Association. Lawrence Rudner of LMP Associates and the Educational Resources Information Center for Tests, Measurement and Evaluation (ERIC/TME), who testified as an expert witness for the parents, told FairTest that he hoped a victory would set an important legal precedent for changing validation procedures in educational tests. He pointed out that federal government requirements for employment test validation are far more rigorous.

The parent plaintiffs raised a series of objections to the validation of the test. First, the plaintiffs argued that use of the test involved a prediction as to the parents' competence to teach at home. Thus, the state should have conducted a predictive validity study to ensure that the test's forecasts were accurate. The plaintiffs demonstrated that each of the eight home-schooled children whose parents failed the EEE scored above the state average on the Comprehensive Tests of Basic Skills (CTBS). According to the plaintiffs, this demonstrated the inaccuracy of the implied prediction that parents who fail the test cannot adequately teach their children.

Second, the plaintiffs alleged that the quality of the validity panel was poor and its procedures faulty. Several panelists, including the dean of the school of education of a state university, acknowledged in court that they were not qualified because they knew little about home schooling. The EEE includes questions on classroom management, discipline, and student motivation, all of which are handled differently in the home-schooling environment. Moreover, parents who had been on the validity panel reported they had been told that if they found all the items on the test irrelevant, their evaluations would be discarded. As a result, they felt forced to approve questions they considered to be of poor quality.

Third, application of the data from the validity study was challenged as faulty. For example, on the math portion of the test, the panel rated 39 items as relevant, but test-takers were required to correctly answer 40 items in order to pass. That is, home schooling parents were required to correctly answer admittedly irrelevant questions in order to be able to teach at home.

The trial was held the week of September 10. Additional legal briefs were due October 15, with the judge's decision to follow. ❐

I'M SORRY. YOUR KID IS A NERD.
BONK!

The Standardized Athletic Aptitude Test

Employment Testing Victories

(Continued from page 1)

Battery (GATB) "because of concerns over whether the test adequately serves all individuals, including minorities, veterans, those with disabilities and older workers."

The GATB had been used by approximately three dozen state and 800 local public employment service offices to make pre-hiring selection and referral decisions, largely for blue-collar jobs. The suspension also applies to similar tests developed by some State Employment Service Agencies and to use of the GATB by Job Training Partnership Act programs across the country.

At the same time, DOL announced that it will conduct a two-year research program to review all its employment aptitude measurement procedures. During this study period, the GATB can only be used for "counselling" job applicants.

The Labor Department action followed a National Research Council (NRC) report which criticized the concept of "validity generalization" on which the GATB is based. Proponents of "validity generalization" believe that one test can predict equally well for all jobs. The NRC also concluded that the accuracy of the GATB had been greatly exaggerated by its promoters. Minorities were shown to be hurt most by the test's inaccurate predictions. As a result, the NRC recommended that "no job seeker be obliged to take the GATB."

DOL's action has been applauded by leading testing reform organizations, including the Mexican American Legal Defense and Educational Fund and the Lawyers Committee for Civil Rights Under Law. However, some advocates are concerned that DOL's two-year research program will be insufficient and that data collected for the study will not be available for public review.

New Federal Civil Service Test

After nine years, exams are once again being administered for professional jobs with the federal government. The previous test, the Professional and Administrative Career Examinations (PACE), was dropped under a consent decree, in response to charges that it discriminated against minorities. Forty-two percent of whites had passed PACE, but only 5% of blacks and 13% of Hispanics. Since positions were filled based on test score rankings, blacks were actually hired for less than 1% of all jobs.

This spring, in the first administration of the new test, the Administrative Careers with America (ACWA), minorities passed at a rate seven times higher than under PACE. Moreover, applicants with high college grades are not required to take the test to obtain government professional jobs. Though some adverse impact still exists under this system, the use of ACWA appears to be a major step forward. Initial indications are that nearly a quarter of new hires will be black or Hispanic. Advocates plan to continue monitoring the results to insure the federal government remains a leader for equal employment opportunity. They will also look for innovations in ACWA that can be applied to other hiring processes.

Integrity Tests Dissected

In a third important federal government initiative, the Congressional Office of Technology Assessment (OTA) has released a report critical of so-called "integrity" or "honesty" tests. An OTA publication released in late September, titled *The Use of Integrity Tests for Pre-Employment Screening*, finds that "existing research is insufficient as a basis for supporting the assertion that these tests can reliably predict dishonest behavior in the workplace."

The OTA report states, "Errors in test results, potential discriminatory impacts, and potential violations of privacy raise important public policy issues pertaining to the use of integrity tests."

Copies of the report are available for $4.00 from: Gov't. Printing Office Washington, DC 20402-9325. ❐

FairTest News

Staff

FairTest welcomes **Laura Pierce**, a recent graduate of Brown University, as our new Administrative Assistant...**Mike Friedman**, formerly in charge of publications and office management, left in August. Mike's skills and presence are missed by everyone. **Jamie Souder** is FairTest's new Office/Publications Manager.

National Policy Panel

FairTest welcomes **Dr. Lucinda Lee Katz**, Director of the University of Chicago Laboratory Schools, to its National Policy Panel.

Upcoming Speaking Engagements

- **November 10**, American Association for the Advancement of Science, Washington, DC.
- **November 16**, National Association for the Education of Young Children, Washington, DC.

THE WALL STREET JOURNAL REPORTS.

© 1990 Dow Jones & Company, Inc. All Rights Reserved.

FRIDAY, FEBRUARY 9, 1990 THE WALL STREET JOURNAL

EDUCATION

A TEST FOR EVERY TASK

No matter what you want to do, your prospective employer may make you pass an exam before you can do it

By Gary Putka

IN THESE TIMES of demanding employers and a work force found wanting, maybe there isn't really a job for everyone after all. But there's bound to be a test to see whether anyone qualifies.

You may be rustic, handy with a wrench, and interested in a job installing wood stoves. But, the contractor asks, have you passed the Wood Heating Alliance test for certified wood-stove workmen?

Think you can cut it framing fancy etchings? To get work from a gallery or dealer, you may need to take the Certified Picture Framers Examination first.

Paper-and-pencil tests of work skills, both professional and basic, are gaining in popularity. Several types of workers are also imposing new tests on themselves — often to meet new regulations, to establish credentials for customers or employers, or in the case of health-care providers, to be paid by insurance companies. Even aerobics instructors have a certification test, launched in 1987.

"That isn't as silly as it sounds," insists Benjamin Shimberg, a testing expert familiar with the aerobics examination. "If you get some jock up there who doesn't realize his class may not have the same aerobic powers that the instructor does, you're going to have a lot of people dropping."

Maybe. But whatever their value, there are more job-related tests to come. The American Business Conference, a Washington-based trade group for medium-sized companies, is working with the Educational Testing Service, publisher of the Scholastic Aptitude Test, to develop a multi-employer assessment for high-school graduates who don't go to college. And the U.S. Supreme Court may have opened the floodgates to more employment testing with a decision last June that raised the burden of proof in job-discrimination suits.

Critics continue to complain — and sue — based on allegations of discrimination against minorities and women in the test questions. Another complaint is that many multiple-choice tests don't discriminate at all — that is, they can't really determine who will do the job well and who will do it poorly.

"There's little evidence that many of these tests really measure what people know how to do," says Robert Schaeffer, a spokesman for the National Center for Fair and Open Testing, a consumers' and minorities' activist group in Cambridge, Mass. "Do you want a typist who can correctly answer a multiple-choice question on what row the 'A' key is found, or do you want one who types 60 words a minute?"

Employers say that tests are often the only economical way to assess minimum competence on a large scale. The growing use of such tests, acknowledges Barry Rogstad, president of the American Business Conference, is an indication that employers no longer trust the schools to tell them whether people are qualified. Mr. Rogstad says the conference's national workers' assessment will strive to measure actual workplace skills such as the abilities to communicate and work as a team, in addition to measuring reading levels and numbers proficiency.

"A kid can't go to college without taking the college-board tests," says Mr. Rogstad. "I want to make it so that if you want to control your employment destiny and get into the work force, you will have to be in this ball game and take this assessment. If we wait around for the perfect form of assessment, we may never get this going." ■

MR. PUTKA COVERS EDUCATION FROM THE WALL STREET JOURNAL'S BOSTON BUREAU.

NAME AN OCCUPATION, and there's probably a test that goes with it. States and trade associations give them to grant licenses to, among others, barbers, massage therapists and travel agents. Schools give them before certifying students. And employers give them to job candidates. Here are some of the questions these tests ask, according to "The Book of Tests" (compiled by Allen D. Bragdon, copyright 1989, published by Harper & Row). ■

BARBER

1. A man with a prominent nose should have a
- a. pyramid mustache
- b. large mustache
- c. narrow mustache
- d. small mustache

2. Lather is rubbed into the beard with the
- a. knuckles
- b. palms
- c. nail tips
- d. cushion tips of fingers

BARTENDER

3. Which one of the following drinks is always shaken?
- a. California Root Beer
- b. Screwdriver
- c. Gibson
- d. Sombrero

4. How many ounces are there in a drink served "on the rocks"?
- a. 2 ounces
- b. 1-1/2 ounces
- c. 1 ounce
- d. 3/4 ounce

BUILDING CONTRACTOR

5. A sack of cement
- a. weighs 94 pounds
- b. weighs 376 pounds
- c. contains 1 cubic foot
- d. A and C

6. What should be painted first when painting a room?
- a. Walls
- b. Windowsills
- c. Trim and mouldings
- d. Ceilings

MASSAGE THERAPIST

7. Nerve strokes are applied
- a. with deep pressure
- b. toward the heart
- c. away from the heart
- d. with circular friction

8. The best type of heat to produce relaxation prior to a massage is
- a. electric
- b. dry
- c. prolonged
- d. moist

BROADCASTER

9. On a telephone call-in program, you should NOT
- a. clarify a topic for a listener
- b. place a time limit on the calls, as this inhibits those persons who wish to express their opinions
- c. express your own opinions on the topics

10. It is always best, when ad-libbing to
- a. stick to one subject
- b. speak at your normal announcing speed
- c. do both of these things

INTERIOR DESIGNER

11. What one element helps to give balance to a room?
- a. A window on every wall
- b. A fireplace
- c. Matching wallpaper and draperies
- d. Furniture of varied heights

12. The best light to read by is
- a. ambient lighting
- b. shielded task lighting
- c. track lighting

FLORAL DESIGNER

13. Which design is not a sympathy design?
- a. Gates ajar
- b. Spray
- c. Cascade bouquet
- d. Cross

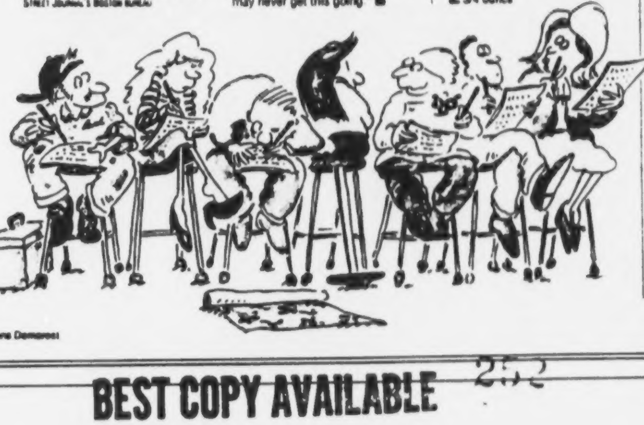
14. Flowers last best if
- a. they are placed in water and preservative
- b. they are placed in water and vodka
- c. they are left out of water
- d. they are put in front of an air conditioner on a hot day

TRAVEL AGENT

15. To obtain the least expensive round-trip air ticket you should
- a. arrive at the airport two hours early
- b. buy two one-way tickets
- c. find three friends to travel with you
- d. book well in advance

16. Which of the following cities is a ski destination?
- a. Davos
- b. Nice
- c. Rome
- d. Dublin

ANSWERS 1b. 2d. 3d. 4b. 5d. 6d. 7c. 8d. 9c. 10c. 11d. 12b. 13c. 14a. 15d. 16a.



Chris Demarest

BEST COPY AVAILABLE

252

New York Guide Released

(Continued from page 16)

state universities to ensure that new teachers understand performance-based methods of assessment and evaluation. (For a copy of the Campaign's proposal, send an SASE to Joan Harris, NYPIRG, 9 Murray St., New York, NY 10007; (212) 349-6460.)

On the evening news after the press conference, Albany's Channel 6 carried the story, then asked for a call-in vote on the question, "Should standardized tests be banned for children before the third grade?" By the end of the newscast, 701 people telephoned to support the ban, while only 313 were opposed.

Guide to be Widely Distributed

The *Guide to Testing Reform* will be used by the Campaign to inform New Yorkers about the flaws and misuses of most current tests. The booklet also discusses appropriate methods of evaluation and concludes with information about how people can participate in the New York Campaign for Genuine Accountability. A Spanish-language edition of the *Guide* and a testing reform videotape are also being prepared.

To help make the *Guide* widely available, it has been priced inexpensively, with substantial discounts for bulk purchases: single copies cost $3; 5 for $10; 10 for $15; 50 for $50; and 100 for $75 (for bulk copies, contact Joan Harris at the address above).

Campaign Accelerates

The press conference, proposal to the Regents, and release of the *Guide* coincide with an acceleration of Campaign activities. This fall, public meetings have been held across the state to organize widespread support. At each meeting, participants learned how to present the issue and then discussed how to build a local coalition, do media work, set up a speakers bureau, and participate in the statewide aspects of the Campaign. Local steering committees are being set up to continue work in each region of the state.

Several additional groups and a growing number of individuals have joined the Campaign. Among the new endorsers are: 100 Black Men, Association for Childhood Education International, Parents Coalition for Education in New York City, People About Changing Education (PACE), Puerto Rican Legal Defense and Education Fund, and Rockland Council for Young Children.

New State Assessments?

Also this fall, Commissioner Sobol held a series of meetings across New York to discuss his proposed New Compact for Learning. As part of the Compact, the Education Department has suggested revamping the state's Pupil Evaluation Program.

Under Sobol's proposal, each student would be evaluated in grades 4, 8 and 12. The new assessments would include portfolios, teacher evaluations, and tests in English, math, science and social studies. Sobol's memorandum on the proposed Compact states that the tests "should measure problem-solving skills as well as factual recall. They should not be limited to multiple-choice, standardized tests."

In a letter to Sobol, the Campaign said, "We applaud your initiative, and would like to offer our support." The Campaign expressed hope that the Compact will facilitate a sharp reduction in the use of multiple-choice achievement testing and the introduction of appropriate, performance-based assessment methods at all grade levels across the state. ❐

Teachers Oppose Testing

In a poll commissioned by the Cincinnati Federation of Teachers, 90% of the polled teachers "agreed" or "strongly agreed" that standardized tests distort curriculum and instruction and take up too much school time. The teachers endorsed using other assessment methods.

Worth Reading

Portfolio News, a new publication from the Portfolio Assessment Clearinghouse, is a valuable source of information for teachers and other school personnel seeking to implement portfolio-based assessments. Each issue contains brief reports from practitioners (including their addresses), discussions of issues of concern, and has a bibliography. (Available from the Clearinghouse, c/o San Dieguito Union High School District, 710 Encinitas Boulevard, Encinitas, CA 92024; quarterly, $25.00/yr.)

New NAEP Board Members

The National Assessment Governing Board (NAGB) has a new chair, Richard A. Boyd, and several new members. Boyd, former state superintendent of education in Mississippi, replaces Chester Finn, who remains on the Board. None of the other members whose term had expired, including vice-chair and Texas legislator Wilhelmina Delco, were re-appointed.

FairTest Publications

______ ***FairTest Examiner***, published quarterly. . . $15.00/year (individuals), $25.00 (institutions)

______ Back Issues of ***FairTest Examiner***, $4.00 each: Vol. ___No. ___(All 13 for $40.00)

______ ***Standing Up to The SAT***, by John Weiss, Barbara Beckwith and Bob Schaeffer (192 pp.) . . . $6.95

______ ***Fallout From the Testing Explosion: How 100 Million Standardized Exams Undermine Equity and Excellence in America's Public Schools***, by Noe Medina and Monty Neill (3rd ed., 77pp.) . . $8.95

______ ***Sex Bias in College Admissions Tests: Why Women Lose Out***, by Phyllis Rosser and FairTest staff, 3rd ed. (60 pp.) $7.95

______ ***The Reign of ETS***, by Allan Nairn and Ralph Nader (550 pp.). . . $30

______ ***Beyond Standardized Tests: Admissions Alternatives That Work***, by Amy Allina with FairTest staff (18 pp.) $5.50

______ ***None of the Above, Behind the Myth of Scholastic Aptitude***, by David Owen (327 pp.) . . . $5.95

______ ***Standardized Tests and Our Children: A Guide to Testing Reform in New York***, by FairTest and NYPIRC (30 pp.)...$3.00 (bulk rates available)

Contributions to FairTest

I agree there's no better way to be part of the testing reform movement than supporting FairTest! Enclosed is my tax-deductible contribution!*

______ **FairTest Associate** ($25.00--includes a free one-year subscription to the *FairTest Examiner*)

______ **FairTest Sponsor** ($50.00--includes *Examiner* subscription and *Standing Up to the SAT*)

______ **FairTest Sustainer** ($100.00 or more--includes above plus a 20% discount on publications)

______ **Other**

Publications Total $ ______

Contributions Total $ ______

Check Total $ ______

Name ______

Address ______

City ______ State ______ ZIP ______

Please make checks payable to FairTest and mail to: FairTest, 342 Broadway, Cambridge MA 02139

---

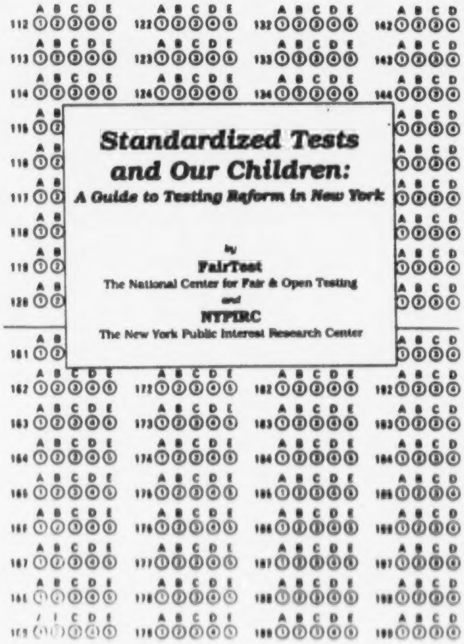
*Contributions to FairTest are fully tax deductible, except for the value of publications received

New Testing Reform Guide Fuels NY Campaign

Standardized Tests and Our Children: A Guide to Testing Reform in New York was released at an October 4 news conference in Albany, NY, held to publicly launch the New York Campaign for Genuine Accountability in Education. The Campaign is an effort, endorsed by over 40 organizations and hundreds of individuals, to replace standardized testing of young children with appropriate forms of assessment.

The 32-page booklet, written for parents and other concerned members of the community, was prepared by FairTest and the New York Public Interest Research Center (NYPIRC), with support from the Aaron Diamond Foundation and the New York Community Trust. (To order a copy, use the form on page 15.)

At the press conference, the Campaign called for a statewide ban on standardized testing in the early childhood grades, to be phased in over a five-year period.



Standardized Tests and Our Children:
A Guide to Testing Reform in New York

by
FairTest
The National Center for Fair & Open Testing
and
NYPIRC
The New York Public Interest Research Center

The Campaign's proposal, submitted to the Board of Regents and Commissioner of Education Thomas Sobol, urged the State Education Department (SED) to work with selected districts, over the next two years, to develop pilot projects in the use of new assessments that are appropriate for young children. As a condition of receiving state aid to implement the new assessments, the pilot districts would be required to immediately stop using current standardized tests.

Joint efforts by the pilot districts and the SED, through the Teacher Centers and other statewide organizations, will enable the development of a trained and experienced staff of experts to help the state's remaining districts implement appropriate assessment prior to full implementation of the testing ban. The Campaign also called for the SED to initiate changes in teacher education at

(Continued on page 14)

FairTest®
National Center for Fair & Open Testing
342 Broadway
Cambridge, MA 02139-1802
(617) 864-4810
(617) 497-2224 (FAX)

Non-Profit
U.S. Postage
PAID
Cambridge, MA
Permit No. 53278

ADDRESS CORRECTION REQUESTED

Larry Roudner
ERIC/TME
3333 K St. NW #200
Washington DC 20009-1275

Moving? Return your label with new address

FairTest
EXAMINER
The National Center for Fair & Open Testing Vol.5, No.1 Winter, 1990-91 ISSN # 0898-2511

FairTest staff members Sarah Stockwell, Cinthia Schuman and Robert Schaeffer at news conference with letter to College Board members responding to "new" SAT.

The "New" SAT:

Less Than Meets the Eye

With great fanfare, the College Board has unveiled "sweeping changes" in the Scholastic Aptitude Test (SAT) which it says will emphasize "critical thinking" and measure students' "problem-solving" and "synthesizing" abilities. But the changes are more public relations puffery than substance.

Instead of introducing reforms addressing the fundamental flaws of the SAT -- its biases, inaccuracy, "coachability," misuses, and lack of relevance to sound college admissions practices -- the College Board has made adjustments which are largely cosmetic.

In reaction, FairTest delivered a letter to College Board members protesting the changes as inadequate. Civil rights, women's, student, and education reform groups, including the Mexican American Legal Defense and Educational Fund (MALDEF), National Organization for Women (NOW), Asian American Legal Defense and Education Fund (AALDEF), Center for Women Policy Studies, United States Student Association (USSA), and Puerto Rican Legal Defense and Education Fund (PRLDEF), joined FairTest in calling on College Board members to reject the revisions.

New Packaging: Same Old Test

The "new improved SAT" differs very little from the current test. For

(Continued on page 12)

FairTest Publishes K-12 Testing Reform Guide

To support the growing national movement to replace standardized, multiple-choice testing with performance-based assessments, FairTest has released *Standardized Tests and Our Children: A Guide to Testing Reform.*

Available in English and Spanish, *Standardized Tests and Our Children* is the first national guide on testing that is easily readable by the average person and also explains how to improve assessment. It is designed to help parents, teachers, administrators and policy-makers, as well as civil rights, education, and community organizations, understand the need to change the way we assess our nation's

(Continued on Page 11)

CONTENTS

No Admissions Test Required 3
Student Loan Rule Challenged 4
NY Campaign Update 6
Groups Oppose Testing ...7
NAEP Standards 11
National Test 16

TM 017 189

New FairTest Board Chair Ron Vera

Vera is New Board Chair

Ron Vera, Esq. was elected to be the new Chair of the Board of Directors after the death of FairTest's former Chair, Denise Carty-Bennia, last fall. A former attorney with the Mexican American Legal Defense and Educational Fund and the firm of Barbosa & Vera, he is now in private practice with the firm of Best, Best & Krieger in Southern California. Vera is also a consultant to the Thomas Rivera Center. In these capacities, he has worked on testing reform litigation and advocacy. ❐

Denise Carty-Bennia Internship Created

To honor Denise Carty-Bennia's life-long commitment to justice, FairTest has created the Denise Carty-Bennia Testing Reform Internship. In addition to a distinguished legal and advocacy career, Denise was Chair of FairTest's Board of Directors and helped found the organization in 1985. She died on September 11, 1990.

Each year, the Denise Carty-Bennia summer Intern will receive funding from FairTest to work on projects addressing racial, ethnic and gender bias in testing.

FairTest is currently seeking Internship applications for this summer from students (preferably graduate students) who have demonstrated a commitment to civil rights, either through academic work or activism. (For more information on how to apply for the Internship, please contact FairTest.)

Friends, colleagues, and admirers of Denise Carty-Bennia and her work are invited to contribute to this ongoing effort by using the form below. ❐

Enclosed is my tax-deductible contribution to the Denise Carty-Bennia Summer Testing Reform Internship:

$25 ___ $50 ___ $100 ___ $250 ___ $500 \_\_\_ Other$ ___

Name ______________________________

Address ______________________________

City ________________ **State** _____ **Zip** ________

Please make checks payable to: FairTest, 342 Broadway, Cambridge, MA 02139

FairTest

National Center for
Fair & Open Testing
342 Broadway
Cambridge, MA 02139
(617) 864-4810 (617) 497-2224 (FAX)
FairTest EXAMINER, Winter 1990-91

Staff:
Cinthia Schuman, **Esq.**, Executive Director
Monty Neill, Ed.D.,Associate Director
Sarah Stockwell, University Tests Director
Veda Wright, Field Organizer
Bob Schaeffer, Public Education Director
Jamie Souder, Publications/Office Mgr.
Laura Pierce, Administrative Assistant

Board of Directors:
Dr. Pamela George, North Carolina Central University; **Dr. Susana Navarro**; **Paul Pottinger**, The Equipment Financing Group; **William Robinson, Esq.**, Dean, D.C. School of Law; **Bob Schaeffer**, Treasurer; **Steve Solomon, Esq.**, Proskauer, Rose, Goetz & Mendelsohn; **Chuck Stone**, *Philadelphia Daily News*; **Ron Vera, Esq.**, Best, Best & Krieger, Board Chair; **John Weiss**, Public Policy Initiatives, Board President

National Policy Panel:
Dr. Leonard Beckum, Duke University; **Dr. Beverly Cole**, NAACP; **Dr. Harold Dent**, Association of Black Psychologists; **Arnold Fege**, National PTA; **Dr. Howard Gardner**, Harvard University; **Dr. Asa Hilliard**, Georgia State University; **Dr. Lucinda Lee Katz**, University of Chicago Laboratory Schools; **U.S. Rep. Matthew Martinez** (D-CA); **Deborah Meier**, Central Park East Secondary School; **Sara Melendez**, University of Bridgeport; **Aurelio Montemayor**, Intercultural Development Research Association; **Harry Pachon**, National Association of Latino Elected and Appointed Officials; **Alfredo de los Santos**, Maricopa Community Colleges; **Eleanor Smeal**, Fund for a Feminist Majority; **Theodore Shaw, Esq.**, NAACP Legal Defense Fund; **Dr. Gail Thomas**, Texas A & M University; **U.S. Rep. Ted Weiss** (D-NY).

Funding for FairTest's work comes from the Golden Rule Corporation, Ford Foundation, Rockefeller Family Fund, Philip M. Stern Family Fund, Aaron Diamond Foundation, Edward Hazen Foundation, Lilly Endowment, New York Community Trust, and many individual contributors.

UNIVERSITY TESTING

"No Admissions Test Required"

The long-term trend toward de-emphasizing college admissions tests has gained momentum in recent months. Bates College in Maine decided to make all test scores optional for admission; the newly established Massachusetts School of Law became the first law school in the nation not to require LSAT scores for admission; and Lewis & Clark College in Oregon added a "portfolio path" option to its application process.

Bates College Drops Test Score Requirement

Last November, after a meticulous five-year study which proved that the SAT adds little predictive value to the high school record, the Bates College faculty voted to make standardized tests completely optional for admission. The faculty vote was virtually unanimous: 84 to 1.

The action builds on Bates' 1984 decision to make the SAT optional while still requiring three College Board Achievement Tests. Last year, in comparing five years of enrollees who submitted SAT scores with those who didn't, Bates found that while "non-submitters" averaged 160 points lower on the SAT (scores were collected after enrollment for research purposes), their freshman grade point average -- which is what SAT scores are supposed to predict -- was only .05 lower than that of "submitters." The academic survival rate of non-submitters was nearly flawless and better than that of submitters: in five years only one of the 14 students dismissed from Bates for academic reasons was a SAT non-submitter.

Though applications have declined at most colleges in recent years, Bates has experienced a substantial increase in the years following its SAT-optional decision. Members of minority groups elected to withhold their SAT scores at a higher rate than the total applicant pool, and applications and enrollment by these students more than doubled during this time.

According to Dean of Admissions and Financial Aid, William Hiss, writing in Bates' September 1990 alumni magazine, Bates' decision to make the SAT optional in 1984 was spurred by concerns that the SAT was restricting the applicant pool Bates wished to attract and that the SAT was having a harmful effect on students:

> *...there was a growing sense that SAT's were sometimes unhelpful, misleading, and unpredictive for students in whom Bates has always been interested: minority students, rural and Maine students, first-generation college students, and foreign-born or bilingual students. Bates' reasonably high average SAT scores were scaring many good students off. More families than we would like use average SAT scores published in guidebooks to decide if Bates is an appropriate institution for their children, and falsely conclude either that Bates is too difficult academically or that their sons or daughters could not be admitted.*

Hiss and others at Bates were also concerned about the overall impact of the SAT on students and the admissions process:

> *Average SAT scores can become a kind of insidious short-hand or code for the health and quality of an institution. In individual cases, low or average scores exert sometimes not very subtle pressure on admissions officers to deny students who we may think for other reasons can do good work at our institution. If SAT's are the volatile barometers of what the public believes to be the quality of an institution, they become a self-fulfilling institutional prophecy.*

Bates' experience demonstrates once again that even very selective schools don't need the SAT--or any test score--in order to choose their freshman classes. At Bates, many non-submitters who turned out to be fine students might never have been admitted -- or would have been discouraged from applying -- without the SAT optional policy.

No LSAT Required at New Law School

The recently established Massachusetts School of Law (MSL), which won the right to grant degrees last summer, has decided not to require the LSAT for admission. In explaining MSL's decision, Lawrence Velvel, Dean of the law school said:

> *In our judgment, there are so many problems with the test, and it is so socially destructive, that no law school should have to defend a decision not to require it -- a decision which mirrors the similar unwillingness of increasing numbers of other kinds of*

(Continued on page 5)

Student Loan Rule Requiring Schools to Produce Scores or Lose Aid is Challenged

A new federal regulation, which requires that all students who do not have high school diplomas be tested before they can enroll in college or post-secondary educational institutions, has been temporarily blocked by a federal judge in California.

Designed to curb student loan defaults and abuses by trade schools, the rule would force institutions to use scores on independently administered examinations to demonstrate that students without high school diplomas have the "ability to benefit" from post-secondary education. Under the regulation, institutions which do not comply will lose their federal funds.

Education and civil rights leaders claim that the regulation goes beyond the law it was intended to implement by applying the testing requirement to applicants who are not seeking financial aid.

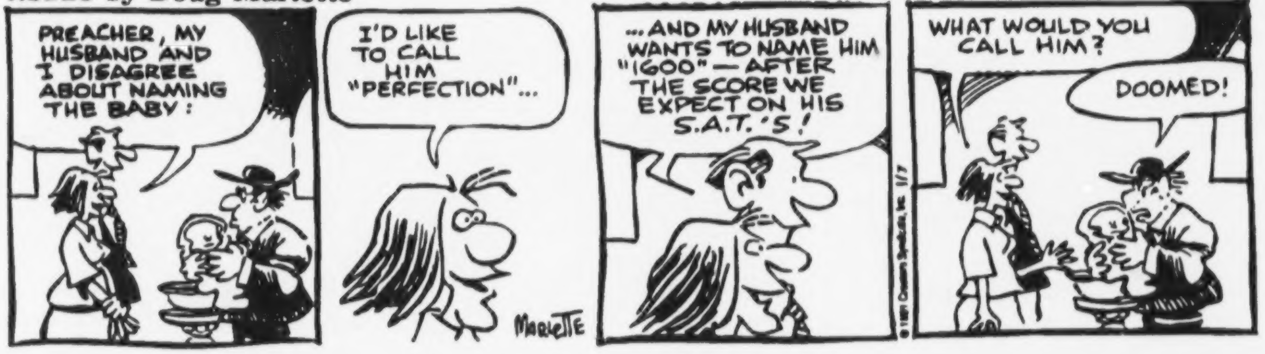
The rule was challenged by the California Community Colleges, which claimed that there was not enough time to comply with the provisions, that the regulation goes beyond the intent of the law, and that the rule contradicts a California law prohibiting the use of a test to deny a student admission. In addition, they charged that the government violated federal laws by not allowing public comment on the regulation before it was issued. If the regulation had been implemented on January 1, 1991 as planned, the California Community Colleges stood to lose up to $283 million in federal aid.

Some community college and trade school officials fear the measure would have the effect of preventing many minority, immigrant and low-income students from obtaining a higher education. Students would be barred from enrolling until they had graduated from high school, obtained a GED, or achieved a minimum cut-score (defined to be at least one standard deviation below the mean) on one of several standardized tests approved by the Department of Education. These approved exams include the SAT and ACT, among others -- tests which, by themselves, are not valid for predicting academic success. Since members of minority groups and those with low incomes tend to score far below average on these standardized tests, it is likely that many of these potential students would be permanently barred from enrolling. Immigrants--some of whom may be highly educated but lack documentation of their education -- may be turned away from English classes solely because they cannot pass a test in English.

A number of higher education groups, including the American Council on Education, Association of Community and Junior Colleges, and the California Community Colleges, have indicated that they intend to press the Education Department to issue new regulations.

In addition to the California suit, a restraining order was issued by a federal judge in response to a complaint brought by E.F. Wonderlic Personnel Test, Inc., an exam publisher. Wonderlic claimed that its exams were unfairly omitted from the list of tests institutions could use. The judge later withdrew the restraining order, after ordering that two Wonderlic tests be added to the list. The Education Department has complied, adding the Wonderlic tests, as well as four others, to the list of approved exams. ❐

KUDZU by Doug Marlette


PREACHER, MY HUSBAND AND I DISAGREE ABOUT NAMING THE BABY:
I'D LIKE TO CALL HIM "PERFECTION"...
...AND MY HUSBAND WANTS TO NAME HIM "1600" — AFTER THE SCORE WE EXPECT ON HIS S.A.T.'S!
WHAT WOULD YOU CALL HIM?
DOOMED!
MARLETTE

LSAT to Change... Again

Only two years after a major overhaul of the Law School Admissions Test, the Law School Admissions Council (LSAC) has announced it will once again significantly revise the format and structure of the test. According to LSAC, the new changes primarily "address problems in the LSAT score distribution, including the number of high scores, the steady climb in average score that has occurred over the past several years, and the continued larger proportion of the test takers earning scores above the mean than below it." Some observers point to the increase in test coaching as a possible explanation for score inflation.

In the new test, scores will be reported in a 61 point range from 120-180 points; the current, narrower, score range is 10-48 points. The structure of the test will also change from three 45-minute sections to four 35-minute sections. There will still be one experimental section on the test.

Critics say the revisions are designed to better enable top law schools to distinguish among high test scorers for admission, but will make little difference to schools that are less selective. ❐

No Admissions Test Required

(Continued from page 3)

schools to require other aptitude tests such as the SAT's or GRE's.

According to Velvel, the school's first graduating class (comprised of 14 transfer students) had a pass rate of 64% on the bar exam, which is "higher than bar passage rates achieved in many states by long established, ABA-approved schools, whose admissions processes place great importance on the LSAT." The LSAT scores of MSL students who passed the bar examination ranged from 16 to 30 (on a 1 - 48 scale), demonstrating that scores commonly considered to be very low do not necessarily indicate an inability to pass the bar examination. Velvel says that these scores are considered so low that "[t]raditional legal educators will frankly admit they would not have believed that people with such scores could pass the bar examination."

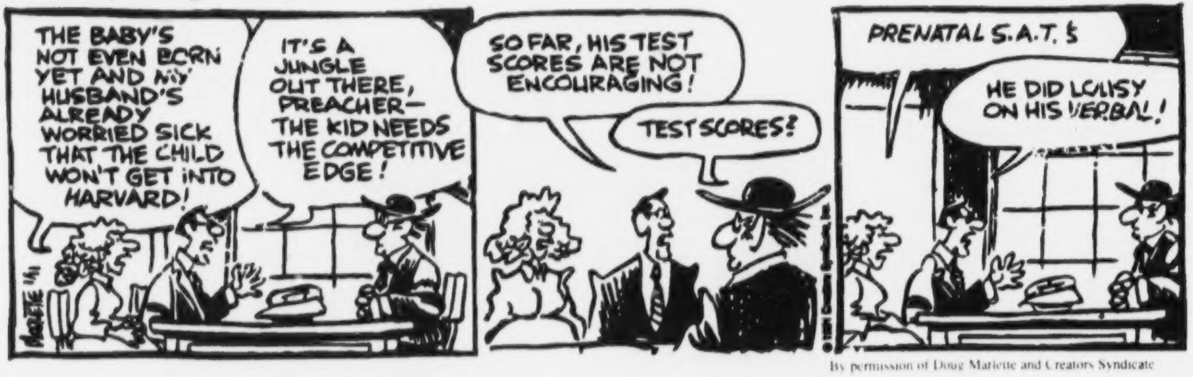
Lewis & Clark's "Portfolio Path"

Lewis & Clark College in Portland, Oregon, added an innovative option to its application process last fall. The school's "Portfolio Path to Admissions" allows students to be judged on the basis of materials which they choose to submit rather than the standard essay and SAT scores. A student's portfolio can include science projects, a mathematical proof, sample work from advanced placement or honors courses, tapes of musical or theatrical events, or photographs of artistic work. Test scores are not required.

Dean of Admissions Peter Brown expects the portfolio option to be popular with students. Even though it was adopted too late to be publicized in college guides or even Lewis & Clark's own "viewbook," nearly 30% of "early decision" applicants submitted portfolios rather than test scores last fall.

These latest actions at Bates College, the Massachusetts School of Law, and Lewis & Clark College make it increasingly clear that not only can institutions of higher learning manage well without multiple-choice admissions tests, but schools and students benefit when tests are made optional. ❐

KUDZU by Doug Marlette


THE BABY'S NOT EVEN BORN YET AND MY HUSBAND'S ALREADY WORRIED SICK THAT THE CHILD WON'T GET INTO HARVARD!
IT'S A JUNGLE OUT THERE, PREACHER— THE KID NEEDS THE COMPETITIVE EDGE!
SO FAR, HIS TEST SCORES ARE NOT ENCOURAGING!
TEST SCORES?
PRENATAL S.A.T.'S
HE DID LOUSY ON HIS VERBAL!
MARLETTE

By permission of Doug Marlette and Creators Syndicate

BEST COPY AVAILABLE

New York Campaign Update

The New York Campaign for Genuine Accountability in Education, a coalition of organizations and individuals working to stop the inappropriate testing of young children, has unveiled a two-part "Proposal to Replace Standardized Tests With Appropriate Assessment in the Early Grades."

The first part of the new proposal calls for a ban on standardized tests in pre-kindergarten through grade 2 by phasing out the exams over a five-year period. The second would create a pilot program for school districts to implement alternative assessments. Participating districts would have to agree to stop using multiple-choice tests in exchange for the funding of authentic, performance-based assessments. The proposal is in two parts so that the ban is not dependent on money being available at a time of tight state budget fiscal constraints.

In other activities, local residents have begun forming steering committees to support the proposal. This follows successful regional meetings in Buffalo, New York City, Syracuse, Albany, Long Island, Mamaroneck, and Rochester.

New endorsers of the Campaign's "Statement to Urge a Ban on Standardized Testing of Young Children" include: Albany Citizens for Education, Buffalo State College Department of Elementary Education Graduate Reading Faculty, District 15 Community School Board of New York City, Early Childhood Research Center, Puerto Rican and Latino Roundtable, Rochester Association for the Education of Young Children (AEYC), and Syracuse AEYC--bringing the number of endorsing organizations to 45. Also, the New York State Association for the Education of Children has become an active member of the Campaign's statewide steering committee.

Spring will see an expansion of Campaign activities through increased public education, outreach, and dissemination of *Standardized Tests and Our Children*, the Campaign's testing reform guide. ❐

NY Rules Against Test Cut-Score Use

New York Commissioner of Education Thomas Sobol has issued an order prohibiting a district's use of test scores as the sole determinant for eligibility for educational enrichment programs. The order states that a district policy:

> *"which denies a student the possibility of further consideration for placement in an enrichment program solely on the student's failure to achieve above a certain score on a subpart of the ITBS [Iowa Test of Basic Skills] is not a legitimate measure for screening a student's capacity for success in an enriched program and is, therefore, arbitrary, capricious and contrary to sound educational policy."*
>
> *(Order #12433, issued on December 7, 1990.)*

Sobol's ruling resulted from a petition brought by the mother of an eighth-grade student who had been excluded from enrichment programs because of her test scores. The plaintiffs were supported by an *amicus curiae* brief from the NAACP Legal Defense and Educational Fund (LDF).

In his decision, Sobol cited the ITBS test manual that states, "Test results should be used to supplement, not to replace, teacher judgement. They should be used in conjunction with everything else the teacher knows about the pupil to meet individual instructional needs." Sobol concluded that the district policy "was inconsistent with the specific guidelines provided" by the testmaker. He ordered the school district to revise its "policy regarding eligibility for enrichment programs" and consider grades and recommendations along with the test scores.

Attorney Norman Chachkin of the LDF stated that the decision "is very significant for districts in New York and, I hope, nationally." The school district has the right to appeal the commissioner's decision in court, but thus far has not done so. ❐

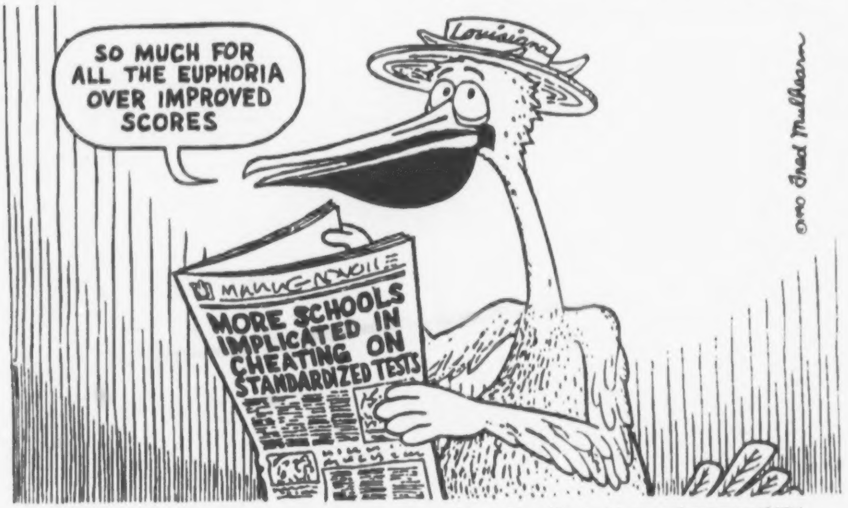
More Groups Oppose Standardized Testing

More national and state organizations have joined the growing chorus of voices against the harmful effects of standardized testing, particularly the testing of young children.

The **American Federation of Teachers** (AFT), at its annual meeting, endorsed a resolution on "Genuine Accountability in Testing." After listing many problems associated with reliance on multiple-choice, standardized tests, the AFT resolved to support alternative assessments that are "based on asking students to demonstrate what we want them to be able to know and do." The AFT further resolved "that a timetable should be set for phasing out standardized, multiple-choice tests and replacing them with reliable, valid and educationally appropriate alternatives." This resolution was in strong agreement with the FairTest-initiated *Statement of Genuine Accountability in Education*, which the AFT endorsed last year.

The **National School Boards Association** (NSBA) released a report urging school boards to discourage the use of standardized tests for screening, tracking or retaining young children. The report also supports "developmentally appropriate" approaches to early childhood education.

The **Louisiana Association on Children Under Six** (LACUS) has issued a statement opposing the testing of young children. It reads, in part: "LACUS calls for, urges, and henceforth supports the total elimination of formalized testing for all young children under the age of eight. Because employment of formal tests for young children is harmful, inappropriate and error-laden, uses of standardized tests ... must be removed from primary and pre-primary grades." This statement follows one from the Southern Association on Children Under Six (SACUS) calling for a ban on standardized testing of young children, issued last spring. ❐


SO MUCH FOR ALL THE EUPHORIA OVER IMPROVED SCORES
Louisiana
MORE SCHOOLS IMPLICATED IN CHEATING ON STANDARDIZED TESTS

Reprinted from Morning Advocate with permission of artist

Good Teachers Dislike Standardized Tests

"(E)xceptional teachers are seriously disenchanted with norm-referenced standardized tests. They complain that instruction is driven adversely by the need to raise test scores." This conclusion comes from interviews with the 115 finalists in the "Thanks to Teachers" national recognition program. Their views on education and testing were reported in *Voices from the Classroom: Exceptional Teachers Speak*, a joint publication of Apple Computer; The National Foundation for the Improvement of Education; The National Alliance of Business; and Group W.

These award-winning teachers, chosen from 20,000 nominees, believe that testing hinders the development of critical thinking skills. They view standardized tests as "a barrier to effective education" and want to "break the shackles that bind them to wasteful modes of instruction." To replace standardized tests, these excellent teachers support "assessment that is based on actual performance such as portfolios of student work."

The Gallup Organization also polled the 231 semifinalists. These results show that of 15 changes suggested by Gallup as possibly necessary to meet the national goals announced by President Bush and the Governors, only one, "more testing," was viewed as "not necessary" by a majority (58%) of those polled. "More educational research" was deemed "not necessary" by 15% of the teachers polled, the next highest percentage in this category (*NEA Today*, 12/90). ❐

A National Test?

(continued from page 16)

but not publicize, individual test scores; this would force NAEP to change its test design to enable individual scoring, currently not possible with its matrix sample structure.

Neither the trial state-by-state comparisons nor studies of their effects, viability, and validity have been completed. If Congress supports NAEP expansion, it will do so in the absence of information it mandated when it authorized the trial comparisons.

Should NAEP be substantially expanded, it could make NAGB, in effect, an unelected national school board, while NAEP exams could come to dictate a national curriculum. Critics also have argued that NAEP has been slow to utilize performance-based assessment methods, so a NAEP-driven national curriculum could become oriented toward multiple-choice testing.

(An Open Letter on NAEP Expansion, signed by FairTest and 75 other organizations and individuals, discusses these concerns. For a copy, send a $.52 SASE to NAEP Letter, FairTest.)

PEPAC: New Tests

A majority of the President's Education Policy Advisory Committee has recommended developing a national test. PEPAC chair, Paul O'Neill, the Chief Executive Officer of ALCOA, called for testing each student in grades 4, 8 and 12 in five subjects. He proposed, as a start, the development of national tests in reading and math for grade 4, to be implemented within two years.

O'Neill said he was not proposing that the test be made by the federal government or that the federal government require all students to take the tests. Who would make them and how they would become national was left undefined.

PEPAC members argued that the plan is necessary in order to measure whether national goals are being achieved. However, some committee members, as well as other testing experts, oppose the idea, saying the US already has the "most over-tested" students in the world. Any tests developed in this process could have the same kinds of harmful effects as an expanded NAEP, if not worse. If developed quickly for administration to all students, they would almost certainly be predominantly multiple-choice, with potentially disastrous effects on curriculum and instruction.

NCEE: An Examination System

By far the most complex and elaborate proposal is one co-authored by the National Center for Education and the Economy (NCEE) and the Learning Research and Development Center (LRDC) of the University of Pittsburgh to create what they term an "examination system" for the US.

In this plan, states or districts would develop performance-based examination systems, to include open-ended (performance) exams, portfolios, and projects. Students would have to pass "a series of examinations" in order to earn a "certificate of initial mastery." The exams could be taken over a number of years and would be used by employers and post-secondary education institutions in hiring and admissions.

By measuring the acquisition of thinking skills and connecting effort in school to rewards in later life, the proponents believe the exams will compel changes in curriculum and in student effort. The result will be, they claim, "an examination system that will improve instruction and achievement in our schools" so as to produce skilled workers for a "high wage" US economy in the 21st century.

The exams would be "syllabus driven." A national "specification of what students should know and be able to do" would be developed by a National Board of Educational Standards in order to set common national standards in each subject area. The various state and local exam systems could then be "calibrated" to model national exams based on the national syllabus and to other state and local exams. Currently, mass performance-based assessment methods, the curricular agreement, and the calibrating or equating methods do not exist.

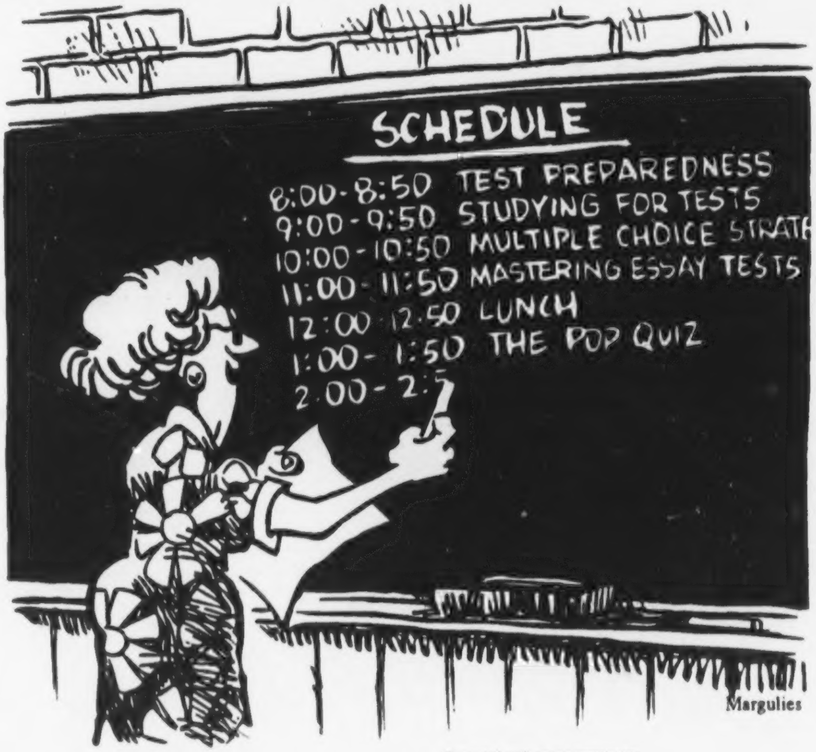
The proposal calls for an initial effort involving 20 or more states and districts. The NCEE and LRDC envision the system being operational for use with individuals -- that is, the awarding of certificates for use by employers and educational institutions -- by the year 2000. The centers have been awarded nearly $2.5 million by the John D. and Catherine T. MacArthur Foundation and the Pew Charitable Trusts to begin work.

Though performance-based assessment is a superior method of evaluation, it is currently being implemented slowly, on a smaller scale, as one part of education reform that includes curriculum change, staff development, teacher empowerment, and school restructuring. Thus, while the plan has many advocates, even supporters have raised critical concerns. These include:

- it will create an unaccountable national school board that, by setting curriculum standards, will lead to a centralized, national education system;
- while assessment can help spur educational improvement, it should not control the direction of educational change;
- the goal of "initial mastery" will encourage sorting and tracking students according to who can best or most quickly reach the goal;
- it does not adequately address equity issues that must be solved for the system to be fair;
- it is too focused on economic needs, to the exclusion of other educational goals, which will then be ignored in school;
- the proposal has seriously underestimated the amount of teacher education and time required to enable teachers to work within the new model;
- the whole process, particularly the calibration, is likely to be too complex, expensive and unwieldy to work;
- when the complexities become clear, the portfolios and projects will, in practice, be reduced to very limited exams, which in turn will have destructive curricular effects, particularly given the centralization and control built into the system; and,
- the net result of this, as with other proposals, will be more, perhaps different, testing, but not real school reform.

SCANS

The Secretary's Commission on Achieving Necessary Skills (SCANS), initiated by former Labor Secretary Elizabeth Dole, is working to define "the skills needed to succeed in the new economy." The commission is charged to "recommend the skills required by high school graduates to achieve work readiness ... suggest the most effective ways to measure individual's abilities ... [and] propose acceptable levels of proficiency." SCANS shares overlapping objectives and personnel with the National Center for Education and the Economy and the Learning Resource Development Center, including the idea of the "certificate of initial mastery." Observers have suggested that the NCEE could become the development arm for SCANS.


SCHEDULE
8:00-8:50 TEST PREPAREDNESS
9:00-9:50 STUDYING FOR TESTS
10:00-10:50 MULTIPLE CHOICE STRATE
11:00-11:50 MASTERING ESSAY TESTS
12:00-12:50 LUNCH
1:00-1:50 THE POP QUIZ
2:00-2:5
Margulies

Reprinted with artist's permission

SCANS has begun to establish its lists of workplace skills. But, critics have argued, it has followed a narrow, outmoded process of breaking down tasks into small component parts. This approach could reinforce types of instruction and assessment that SCANS professes to want to change. Some educators are alarmed that SCANS could become highly influential in national testing through federal involvement in vocational education and its connection to NCEE/LRDC. (SCANS documents are available from the Dept. of Labor at (202) 523-4840.)

State Consortia

Both the Education Commission of the States (ECS) and the Council of Chief State School Officers (CCSSO) have separately begun organizing consortia of states and districts interested in performance-based assessments. These assess-

(Continued on page 10)

A National Test

(Continued from page 9)

ments can be expensive and time-consuming to develop, so state leaders thought collaboration could be useful. Some money to help the process may be available from the Office of Educational Research and Instruction (OERI), which is scheduled for Congressional reauthorization this year.

The Chiefs' "Interstate Consortium on Alternative Methods of Student Assessment" includes a "laboratory for developing alternative methods" that will "assist states in designing, developing, and testing alternative assessment methods," with a focus on solving "the most pressing and difficult technical issues," and an "information exchange on innovations in student assessment."

ECS is planning a conference that will enable assessment developers and policy leaders at the state level to discuss the development of new assessments and encourage the continued exchange of knowledge and experience among the states. The two organizations have begun to collaborate so as to avoid duplication of effort and enhance the process.

Measuring National Goals

Currently, determining how to attain national goals and measuring progress toward them is being led by a panel comprised of governors and members of the Bush Administration. Panel committees, including one on assessment, are due to report in early February. The panel may develop national testing recommendations. Sen. Jeff Bingaman (D-NM) last year proposed that a national panel of leading educators monitor the nation's educational progress. In passing the Senate, the legislation was amended to create an 18-member "National Council on Educational Goals." The bill was opposed by the Bush Administration and not included in the final education legislation of the last Congress. It is expected to be re-introduced, perhaps in modified form, in 1991.

The NGA/Administration panel, which has been holding hearings on educational goals, has no legal authority, whereas the National Council would. The relationship between the two, or even whether the first would continue to exist, is not clear; nor is the relationship between the proposed Council and NAGB. In short, the process of monitoring national educational progress and the issue of who controls the monitoring, remains to be decided this year.

In addition to all these proposals, the National Indicators Panel, set up by OERI, is due to report this spring. Included in its report will be recommendations on testing and assessment as part of national indicators. Other members of Congress are reportedly considering hearings or legislation, and various business groups have offered suggestions for national testing.

Conclusion

Decisions made this year and next could fundamentally affect the shape of assessment at the national level, and with that, the direction of education over the next decade. Discussion of testing at the federal level has never been so intensive. A number of education lobbyists believe that some kind of national testing is almost certain to pass this year or next. Others suggest that the complexity of the issue and the ramifications of any action may cause Congress to not endorse a national test, at least this year. That would still leave the proposals by PEPAC and the NCEE/LRDC in development.

FairTest will be actively involved, working with other organizations to promote a reasoned and useful role for the federal government in monitoring educational progress and supporting helpful educational change. These developments, as well as continuing information on national testing efforts, will be discussed in future *Examiner* articles. ☐

Boston Drops MAT

The Boston School Committee has voted to drop the requirement that students must attain certain scores on the Metropolitan Achievement Test (MAT) in order to graduate. As part of restructuring the city's high schools, the School Committee also voted that students should demonstrate competencies rather than accumulate course credits in order to graduate. ☐

Great City Schools Support Alternatives

A task force initiated by the Council of Great City Schools, including 47 urban school districts and 70 national education, civil rights, human service, and business organizations, has called for the use of alternatives to standardized tests to assess student academic performance. The recommendation was part of six goals designed to help reform urban education. ☐

NAEP Standards and Reports

The National Assessment Governing Board (NAGB) has voted to delay, for three months, release of a report on the 1990 National Assessment of Educational Progress (NAEP) math assessment. Controversy surrounding NAGB's proposal to set achievement standards for students to meet on the math assessment caused the delay. The NAEP report, the first to include state-by-state comparisons, is now scheduled to be released in September 1991. The Board will decide later whether to include the standards in the report.

The standard-setting process for the math assessment was itself slowed as the panels charged with the task debated what test items should represent "basic," "proficient," and "advanced" levels of math achievement at grades 4, 8 and 12. The standards were finally made public in mid-November.

Later in November, NAGB heard from opponents and supporters of the standard-setting. Frank Betts of the Association for Curriculum and Development (ASCD) charged, "There is very real potential for harm in reporting three levels of achievement" based on comparing groups of students who exhibit very small differences in test performance. Herbert Rosenthal from the Coalition of Essential Schools warned that too many of the test questions were based on a wrong view of math education and would hinder needed reforms in math instruction (*Education Week*, 12/5/90).

At a January 8, 1991, NAGB forum, Greg Anrig of ETS repeated the charge, raised to NAGB over the previous year by many critics, that constructing standards from an already existing, mostly multiple-choice test was a backwards approach and would not help needed educational reform. Anrig claimed that the process used a test that measured "what is" to try to measure "what ought to be." He and others urged postponement of the use of standards. ETS makes the NAEP exams.

Board members and supporters defended the standard-setting, arguing that because it only samples students, it is not likely to affect school curricula, while the standards would make NAEP data more useful to policymakers.

Poor Performance on NAEP

In a summary of its assessments from 1986 and 1988, a NAEP report concluded that educational performance in reading, math, writing, history, science, civics and geography has shown little improvement in 20 years, gaps between whites and minorities remain "unacceptably large," and gender differences remain unchanged. *America's Challenge* noted some gains by minorities, but a decrease in the overall percentage of students who could solve more complex problems in various subject areas. The report also documented the persistence of poor teaching methods, even in advanced classes, and the low numbers of students exposed to challenging academic work.

(*America's Challenge* is available from NAEP, P.O. Box 6710, Princeton, NJ 08541; $12. The report on NAEP math standards is free from NAGB, 1100 L St., NW, #7322, Washington, DC 20005.) ☐

Author's Query

For a book on the impact of standardized testing on young children, parents, and teachers, Samuel J. Meisels would appreciate hearing from anyone willing to share anecdotes and other information (newspaper articles, unpublished papers, personal recollections, etc.) regarding the use and abuse of testing in pre-K through grade 3.

Write to:
Samuel J. Meisels
School of Education
The University of Michigan
Ann Arbor, MI 48109

Reform Guide

(continued from page 1)

students and begin to make the changes.

The 32-page pamphlet features chapters on "How Standardized Tests are Used"; "What Are Standardized Tests?"; "What's Wrong with Standardized Tests?"; "Better Ways to Evaluate Students"; and "What You Can Do." It also contains special sections on parents' rights, the history of testing, commonly used tests, and what test scores mean, as well as appendices on useful resources and testing terms.

To help make the Guide widely available to groups active in testing reform, it has been priced inexpensively, with substantial discounts for bulk purchases: single copies cost $3; 5 for $10; 10 for $15; 50 for $50; and 100 for $75. Bulk rates also apply on orders combining English and Spanish versions. Use the order form on page 15. ☐

"New" SAT: Less than Meets the Eye

(Continued from page 1)

the most part, sections of the current test have simply been shuffled and renamed.

The "SAT-I: Reasoning" test is essentially the same SAT taken by students for fifty-five years. It will still be predominately multiple-choice, administered in about three hours on a Saturday morning, and reported in two scores: Verbal and Math. Unlike the current test, calculator use will be allowed on the SAT-Math, which will include 10-15 "open-ended" math questions for which students will have to produce their own answers. The SAT-Verbal will include more and longer reading passages, while the antonym item type will be dropped.

The "SAT-II: Subjects" exams are little more than the current Achievement Tests with a different name and two new Asian language tests. Each will still take one hour, and all but one are entirely multiple-choice.

The much-promoted "SAT-Writing" test will be similar to the current English Composition Achievement Test (with essay) with some test questions from the Test of Standard Written English (TSWE), but administered five times a year instead of only once. It will *not* be part of the basic SAT, but rather will be optional (one of the "SAT-II Subject Tests"), just as it is now. At least 40 minutes of this one-hour "writing" test will consist of multiple-choice questions.

Changes Lack Substance

Many critics argue that these changes add little more than a new veneer to a test which needed a comprehensive overhaul. For example, some educators question the value of the SAT writing section, asking how any twenty-minute essay can measure real writing ability, especially for students whose first

The "All New SAT": Myth or Reality?

CURRENT SAT	"NEW' SAT
All multiple-choice questions (145)	About 90% multiple-choice (115-130 of 125-145)
2 hours 30 minutes long	Longer test
25 reading comprehension questions	38-40 reading comprehension questions
Essay test optional (English Achievement Test)	Essay test optional (SAT II-Writing Test)
Calculators prohibited	Calculators permitted

Many Texas Students Fail New StateTest

About one-half the students in grades 5, 7 and 9 failed the new Texas Assessment of Academic Skills (TAAS), while one-third of the students in grades 3 and 11 failed.

The state claims the mandatory test includes critical-thinking and problem-solving items, making it different from the exam it replaced. The exam is all multiple-choice, except for one writing portion.

On previous Texas exams, Latino and African-American pass rates have been far lower than pass rates for whites. The Texas Education Agency has not yet released demographic information on the TAAS.

Next year, the passing score will increase from 60 to 70. A student must pass the grade 11 test in order to graduate. Students who do not pass it have unlimited opportunities to take the exam again, and remediation is supposed to be provided to assist those students. ❐

Teacher Competency

The American Federation of Teachers, the National Council on Measurement in Education, and the National Education Association have issued *Standards for Teacher Competence in Educational Assessment of Students*. Three years in the making, this set of guidelines is designed to help teacher educators and teachers in developing, selecting, administering, scoring, using, and communicating various assessment methods from norm-referenced, multiple-choice tests to performance-based, in-class assessments. The *Standards* were printed in the Winter 1990 *Educational Measurement*; copies are available from all three organizations. ❐

language is not English and thus may need extra time to polish their work. Others wonder what admissions officers learn from a score on a hastily-written first draft that will be graded in less than a minute.

Nor will adding more and longer reading passages improve the test. Far from making the SAT a better measure of reading and higher order thinking skills, this change will make the test even more coachable. Test preparation firms already say they will train students to go straight to the less time-consuming short-answer questions (such as analogies) first and to read the reading comprehension *questions* before skimming the passage for answers. Research has shown that students can answer nearly 40% of these questions correctly without ever bothering to read the passage.

Despite a great deal of publicity about the addition of so-called "open-ended math questions," non-multiple-choice items will account for no more than 10-15 items out of about 145. And it's not clear that these few items will assess "problem-solving ability." While these questions will measure whether the student correctly gridded-in the right answer, they will reveal nothing about *how* students arrived at their answers, the real goal of assessments purporting to test "problem-solving" ability. Nor can most single-right-answer items assess the ability to solve complex problems. Mechanical errors by students who grid-in the wrong bubbles will be substantial, particularly for test-takers with learning disabilities.

Allowing students to use calculators will add a new bias to the test: students who own sophisticated devices or whose schools can afford to provide them are likely to benefit.

What do Critics Want?

In its three-year, multi-million dollar quest to repackage the SAT, the College Board never addressed the most basic questions surrounding the test:

- **Will the new test be less biased than the current test?**

The College Board's own research proves that the format, context and language of the SAT (none of which will change substantially on the new test) are biased against women and members of minority groups.

- **Will the new test be more accurate than the current test?**

According to College Board data, the SAT's margin of error is 67 points. And two test-takers' scores must differ by at least 142 points in order to show that their abilities really differ. Will the new test provide test users with a score that is not so inexact as to be of questionable value?

- **Will the new test be less coachable?**

A number of studies show score increases of over 100 points with good test coaching programs. Coaching adds another income-related bias to the already existing biases of the SAT. A young African American woman from a family that cannot afford to send her to a coaching course is triply discriminated against by the test on account of her gender, race and inability to pay for coaching. Despite the test-makers' efforts to make the SAT less coachable, test preparation companies expect to have no difficulty gearing up for the changes.

- **If the new test won't perform better than the old test, what's the point?**

The test-makers have admitted that the new test will not predict first year college grades -- the SAT's sole scientific purpose -- better than the current test. Thus, high school grades will continue to predict college grades better than the SAT, according to the College Board's own experts.

- **When will the College Board act responsibly and stop the pervasive and long-standing misuse of its own test?**

National Merit Scholarship semi-finalists are still chosen solely on the basis of Preliminary Scholastic Aptitude Test/National Merit Scholarship Qualifying Test (PSAT/NMSQT) scores, resulting in nearly two-thirds of all scholarships going to males and a minuscule number to minorities. The College Board's own *Handbook* lists nearly 200 institutions which use cut-off scores on the SAT, despite the College Board's guidelines against them. Nor has the Board taken any effective measures to stop the National Collegiate Athletic Association's "Prop 48" regulation which denies athletic eligibility to students who score below certain SAT scores.

A New Direction

College admissions testing now costs the nation over $100 million annually for coaching courses, test preparation materials, and the exams themselves. Instead of squandering millions more implementing minor changes, the College Board should stop promoting the SAT. This money could better be devoted to genuine educational improvement. ☐

(Related story page 14)

The New SAT: Flawed Process, Flawed Product

Internal College Board documents made available to FairTest show that proposed revisions to the SAT were highly controversial within the committee charged with exploring changes in the test.

For example, several members of that committee, known as the "New Possibilities Commission," were critical of a College Board draft report explaining possible changes. One wrote that it was "transparently self-serving, attempting to wrap the SAT in a fabric of noble concern for under-represented minority students." Another memo said, "The title *In Support of All Students* is ironic in view of the text's cavalier dismissal of minorities/women's concerns," and claimed the introduction of the draft had "the tone of a College Board/ETS public relations document." Even incoming College Board Chair Alice Cox concluded in one letter that the draft Commission report by the College Board was "to put it mildly, unsatisfactory." She added that "the draft report lacked substance and did not represent the deliberations of the Commission."

There was virtually no public input into the College Board's decisions, despite the fact that over a million and a half students must take the SAT each year. The only exception was a public hearing held in October of last year by State Senator Art Torres of California, which was organized with the assistance of FairTest. At that hearing, FairTest and representatives from education, civil rights and student groups criticized both the substance of possible revisions and the secret process for developing them.

Unsolicited public commentary did have one clear effect on the College Board's decision-making process, however. Criticism from California Assembly Speaker Willie Brown, Jr., among others, that inclusion of a mandatory essay would be unfair to students whose first language was not English, reportedly led the Board to drop the proposal.

Had the College Board actively sought public input from the constituencies affected by the SAT, the revisions might have had more substance. ❐

Senator Art Torres of California questions "new " SAT.

SUNY Geneseo Finds H.S. Grades Predict Six Times Better Than SAT

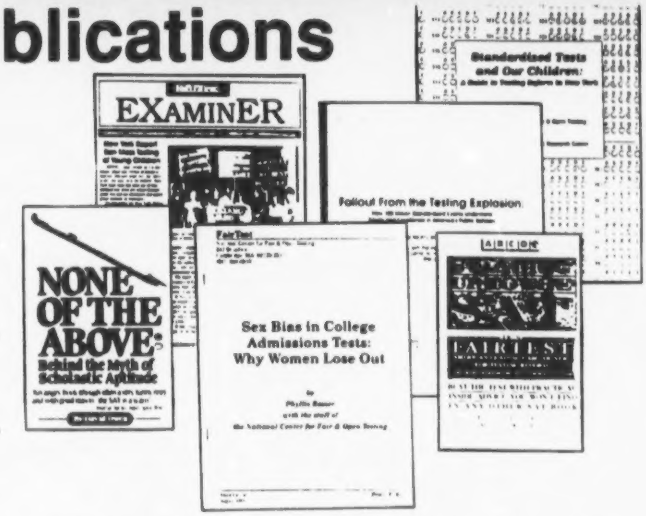
Data from the State University of New York at Geneseo show that high school grades predict first-year college grades at least six times better than SAT-Verbal scores, and 40 times better than SAT-Math. The 1988 study determined an optimum formula for predicting first-year college grades based on data from the class that enrolled two years earlier. It concluded that grades should be weighted much more heavily than the SAT-Verbal, and that the "SAT-Math contributes very little to the prediction" of freshman grades. ❐

GRE Optional at Harvard Divinity School

Students applying for admission to the Harvard Divinity School's Master of Divinity and Master of Theological Studies programs are not required to submit Graduate Record Examination scores. According to school officials, the decision was made over ten years ago when members of the admissions committee concluded that test scores did not accurately reflect the traits they were seeking in candidates. Instead, the committee emphasizes recommendations, college transcripts, and each applicant's personal statement.

This means that two Harvard University graduate programs do not ask applicants to submit test scores. In 1985 the Harvard Business School elected to drop its requirement for Graduate Management Admissions Test scores. At that time John Lynch, Dean of Admissions, said "There are a lot of things that are relevant to admitting MBA candidates, and one of them is not the GMAT." ❐

FairTest Publications



Standardized Tests and Our Children:
EXAMINER
Fallout From the Testing Explosion
NONE OF THE ABOVE: Behind the Myth of Scholastic Aptitude
Sex Bias in College Admissions Tests: Why Women Lose Out
Phyllis Rosser
FAIRTEST

FairTest Examiner, published quarterly...$15.00/year (individuals), $25.00/year (institutions) $____

Back issues of *FairTest Examiner*, $4.00 each:
Vol ____ No ____(All 14 for $40.00) $____

Standing Up to The SAT, by John Weiss, Barbara Beckwith and Bob Schaeffer (192 pp.)...$6.95 $____

Fallout From the Testing Explosion: How 100 Million Standardized Exams Undermine Equity and Excellence in America's Public Schools, by Noe Medina and Monty Neill (45 pp.)...$8.95 $____

Sex Bias in College Admissions Tests: Why Women Lose Out, by Phyllis Rosser and FairTest staff, 3rd ed. (60 pp.)...$7.95 $____

The Reign of ETS, by Allan Nairn and Ralph Nader (550 pp.)...$30.00 $____

Beyond Standardized Tests: Admissions Alternatives That Work, by Amy Allina with FairTest staff (18 pp.)...$5.50 $____

None of the Above, Behind the Myth of Scholastic Aptitude, by David Owen (327 pp.)...$5.95 $____

Standardized Tests and Our Children: A Guide to Testing Reform, by FairTest (32 pp.)... $3; 5 for $10; 10 for $15; 50 for $50; 100 for $75. English and Spanish editions can be combined for bulk prices.

copies in English #____
copies in Spanish #____
Total $____

Standardized Tests and Our Children: A Guide to Testing Reform in New York, by FairTest & NYPIRC (32 pp.)...same prices as above. *Single* copies from FairTest. Bulk copies available from: NYPIRC, 184 Washington Ave., Albany, NY 12210

English #____
Spanish #____
Total $____

Contributions to FairTest

I agree there's no better way to be part of the testing reform movement than supporting FairTest! Enclosed is my tax-deductible contribution!*

☐ **FairTest Associate** ($25.00--includes a free subscription to the *FairTest Examiner*)

☐ **FairTest Sponsor** ($50.00--includes *Examiner* subscription and a FairTest publication of your choice, excluding *Reign of ETS*)

☐ **FairTest Sustainer** ($100.00 or more--includes above, plus a 20% discount on publications)

☐ **Other** ____

Publications Total $____
Contributions Total $____
Check Total $____

270

Name ____
Address ____
City ____ State ____ Zip ____

Please make checks payable to FairTest and mail to: FairTest, 342 Broadway, Cambridge, MA 02139

*Contributions to FairTest are fully tax deductible, except for the value of publications received.

A National Test?

After nearly a decade of continuing debate over school reform, the possibility of a national test of US schoolchildren is now under intense discussion in Washington. While proposals for extensive national testing appear to be moving quickly, the implications for school reform, curriculum improvement, equity, and democratic control over education have only begun to be discussed.

Current testing proposals are, in large part, a response to the National Governors Association (NGA)/Bush Administration national education goals. Despite a tradition of local and state control over education, public support for national standards, curriculum, and assessment has grown, fueled by continuing criticism of schooling, fears of economic decline, and awareness that current textbook and testing practices have tended to create a de facto, but poor quality, national curriculum.

Among the major current proposals:

- The National Assessment of Educational Progress (NAEP) may be expanded to more closely fill the role of a national test.
- The President's Educational Policy Advisory Committee has recommended a national test.
- A plan from the National Center for Education and the Economy has received start-up funding to develop a national, performance-based examination system.
- The Department of Labor Secretary's Commission on Achieving Needed Skills is defining workplace skills and means to assess their acquisition.

And in related matters:

- The Council of Chief State School Officers and the Education Commission of the States are working to form consortia of states to develop new assessments.
- A proposal to create an independent panel to measure progress toward national educational goals, which failed to pass as Congress recessed, is likely to be refiled.

NAEP Expansion

NAEP, "the nation's report card," is a series of tests and surveys given to a sample of students across the nation in different subject areas every two years. In 1987, NAEP was authorized to conduct state-by-state comparisons, on a trial basis, with any states that volunteered, and an independent National Assessment Governing Board (NAGB) was established to govern NAEP.

Expansion of NAEP, supported by the Bush Administration, is almost certain to be considered in 1991. NAGB developed expansion ideas in December 1989 and made similar recommendations to the Department of Education in January 1991. Among them are:

- making voluntary state-by-state comparisons part of each NAEP assessment, with full federal funding for state participation;
- testing more frequently, so that reading and math would be tested every other year, with some NAEP exams given every year; and,
- removing the prohibition on school and district use of NAEP items and tests so they could obtain,

(Continued on Page 8)

FairTest®
National Center for Fair & Open Testing
342 Broadway
Cambridge, MA 02139-1802
(617) 864-4810
(617) 497-2224 (FAX)

10

Non-Profit
U.S. Postage
PAID
Cambridge, MA
Permit No. 53278

ADDRESS CORRECTION REQUESTED

Larry Roudner
ERIC/TME
3333 K St. NW #200
Washington DC 20009-1275

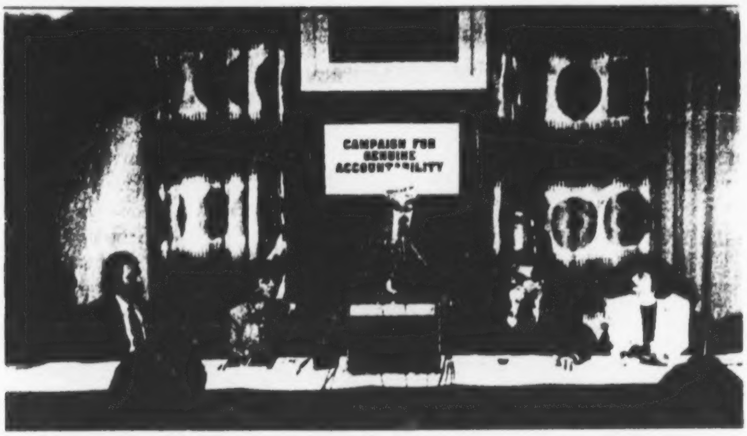
***Do you have an asterisk on your label?** If you do, this is your last issue, so **renew now!**

Moving? Return your label with new address

FairTest EXAMINER

The National Center for Fair & Open Testing Vol. 5, No. 2 Spring, 1991 ISSN # 0898-2511

Groups Oppose National Test



Speakers at news conference on National Testing. (l. to r.) Monty Neill, FairTest; Beverly Cole, NAACP; Ed Keller, National Association of Elementary School Principals; Arnold Fege, National PTA; Sue Bredekamp, National Association for the Education of Young Children

Leaders of more than two dozen major civil rights, education, parent and advocacy organizations issued a strong statement in opposition to national testing at a March 27 press conference in Washington, D.C. Responding to numerous proposals calling for national testing, the Campaign for Genuine Accountability in Education, organized by FairTest, urged "the Bush Administration and the Congress to support education reform by not implementing a national exam at this time."

"There should be no national test unless testmakers and proponents can demonstrate how it will improve the quality of education and provide equity," explained Beverly Cole, Education Director of the NAACP. "We need to stop destructive testing proposals and promote practices that will actually improve education."

"Imposing a national test on every student in America, a current fad in the movement to improve education, will not make students learn more or improve our schools," added Samuel G. Sava, Executive Director of the National Association of Elementary School Principals.

FairTest's Monty Neill told the assembled press, "To ensure improved education for all, we need to reform curriculum and instruction and reshape school governance and structures."

(Continued on page 11)

Test Reformers Win Truth-in-Testing Legal Round

FairTest and its allies have won an important legal victory against testmaker efforts to keep their exams and other testing data secret. On March 12 the Second Circuit Court of Appeals reversed a District Court decision that would have allowed testmakers to administer exams without making the questions and answers available to students and independent researchers after the test (see *Examiner*, Vol. 4, No. 1). In sending the case back for trial, a three-judge panel of the Court of Appeals gave new life to New York's landmark Truth-in-Testing law.

(Continued on page 14)

CONTENTS

New Board Members 2
National Merit 3
College Board's SAT Claims Discredited 5
National Test 6
Hearings on Nat'l. Test......8
Future of NAEP 9
International Testing 10
Ability to Benefit..............16

New FairTest Board Members

FairTest is pleased to welcome three new distinguished members of the Board of Directors: Vito Perrone, Deborah Meier, and Ron Ellis.

Deborah Meier

Deborah Meier is the principal of Central Park East Secondary School. She was chosen as a 1988 MacArthur Fellow in recognition of her innovative leadership at this highly sucessful alternative high school in East Harlem. She has been a long-time testing reform advocate, particularly in the area of reading tests. Meier also serves on the National Board for Professional Teaching Standards, the board of directors of Advocates for Children, and is president of the Center for Collaborative Education.

Ron Ellis is the Director of the Fair Employment Program of the NAACP Legal Defense and Educational Fund. He has extensive experience in anti-discrimination law, including cases involving employment testing, and has authored a number of articles on civil rights and affirmative action. Ellis is an adjunct professor at New York University School of Law, and serves on the board of the Root-Tilden-Snow Scholarship Program.

Ron Ellis

Vito Perrone

Vito Perrone is Senior Lecturer, Director of the Teacher Education Programs, and chair of the Teaching, Curriculum, and Learning Environments Program at Harvard Graduate School of Education. In addition, he is a senior fellow at the Carnegie Foundation for the Advancement of Teaching. He has written extensively about testing reform, most recently authoring "On Standardized Testing," a position paper of the Association for Childhood Education International.

FairTest

National Center for
Fair & Open Testing
342 Broadway
Cambridge, MA 02139-1802
(617) 864-4810 (617) 497-2224 (FAX)
FairTest EXAMINER, Spring 1991

Staff:
Cinthia Schuman, Esq., Executive Director
Monty Neill, Ed.D, Associate Director
Sarah Stockwell, University Tests Director
Veda Wright, Field Organizer
Bob Schaeffer, Public Education Director
Jamie Souder, Publications/Office Mgr.
Laura Pierce, Administrative Assistant

Board of Directors:
Ron Ellis, Esq., NAACP Legal Defense and Educational Fund; **Dr. Pamela George,** North Carolina Central University; **Deborah Meier,** Central Park East Secondary School; **Dr. Susana Navarro; Vito Perrone,** Harvard Graduate School of Education; **William Robinson,** Esq., D.C. School of Law; **Bob Schaeffer,** Treasurer; **Steve Solomon,** Esq., Proskauer, Rose, Goetz & Mendelsohn; **Chuck Stone,** *Philadelphia Daily News;* **Ron Vera,** Esq., Best, Best & Krieger; **John Weiss,** Public Policy Initiatives, President.

National Policy Panel:
Dr. Leonard Beckum, Duke University; **Dr. Beverly Cole,** NAACP; **Dr. Harold Dent,** Association of Black Psychologists; **Arnold Fege,** National PTA; **Dr. Howard Gardner,** Harvard University; **Dr. Asa Hilliard,** Georgia State University; **Dr. Lucinda Lee Katz,** University of Chicago Laboratory Schools; **U.S. Rep. Matthew Martinez** (D-CA); **Sara Melendez,** University of Bridgeport; **Aurelio Montemayor,** Intercultural Development Research Association; **Harry Pachon,** National Association of Latino Elected and Appointed Officials; **Alfredo de los Santos,** Maricopa Community Colleges; **Eleanor Smeal,** Fund for a Feminist Majority; **Theodore Shaw,** Esq., University of Michigan; **Dr. Gail Thomas,** Texas A & M University; **U.S. Rep. Ted Weiss** (D-NY).

Funding for FairTest's work comes from **the Golden Rule Corporation, the Ford Foundation, the Rockefeller Family Fund, the Philip M. Stern Family Fund, the Aaron Diamond Foundation, the Edward Hazen Foundation, the Lilly Endowment, the New York Community Trust, and many individual contributors.**

BEST COPY AVAILABLE

Males Will Receive Three-Fifths of 1991 National Merit Scholarships

Use of Biased PSAT Costs Women and Minorities Millions of Dollars

Three out of every five National Merit Scholarships awarded to high school seniors this spring will go to boys even though girls earn better grades in high school and college, according to a new FairTest analysis. This long-standing disparity is due to reliance on a biased standardized exam, the PSAT/NMSQT (which is derived from the SAT) as the sole criterion for National Merit semifinalist status.

FairTest's state-by-state analysis of the more than 15,000 students eligible for this year's prestigious awards reveals that only 37% of 1990-1991 National Merit Semifinalists are female. Males make up more than 59% of the semifinalist pool; the gender of the remaining students could not be determined from their names.

Young women, as well as members of minority groups, are denied millions of dollars in college aid each year because of National Merit's continued use of a biased test. For example, in New York, the majority of valedictorians are females, yet more than 57% of National Merit semifinalists this year were male. African Americans, Hispanics, Native Americans, and students from recent Asian immigrant groups are also denied their fair share of these awards because of an unfair and irrelevant selection device.

Though the National Merit Scholarship Corporation (NMSC) has confirmed FairTest's reports that the vast majority of winners in past years were male, NMSC has refused to make available any data on the race or ethnicity of Merit Scholars. Since most minority groups have much lower average scores than whites on the SAT, their chances of receiving scholarships are small.

GUIDANCE COUNSELOR

"Great! We're just beginning to learn our multiplication tables, and she says we need to start thinking about the SAT."

Reprinted with artist's permission

Rather than selecting from among the most talented college-bound students, NMSC picks finalists from a pool of outstanding test-takers: Merit Scholarship Semifinalist status is based solely on test scores from the Preliminary Scholastic Aptitude Test/ National Merit Scholarship Qualifying Test (PSAT/NMSQT). The exam, produced by the Educational Testing Service for the College Entrance Examination Board and National Merit, consists of previously-administered, multiple-choice SAT questions. More than 99% of all applicants, including many women and minorities from the very "tops" of their high school classes, are eliminated solely on the basis of their scores on this discriminatory test.

In 1989, a Federal District Court struck down similar scholarship programs in New York -- the Regents and Empire Scholarships -- which awarded scholarships solely on the basis of SAT scores. In that case, Judge John Walker concluded, " *The probability, absent discriminatory causes, that women would consistently score 60 points less on the SAT than men is nearly zero.*"

FairTest and its testing reform allies argue that better ways to award merit scholarships already exist. For example, the Citizens' Scholarship Foundation of America in St. Peter, Minnesota, administers many scholarship programs for communities, corporations and other organizations, considering factors such as high school record, honors courses taken, and leadership, rather than test scores (see *Examiner*, Vol. 4, No. 1). Since research conducted by the Educational Testing Service and the College Board demonstrates that high school grades predict college performance more accurately than test scores, National Merit's continuing use of a discriminatory formula for selecting scholarship winners cannot be justified.

▼ Send stamped, self-addressed envelope to FairTest for state-by-state breakdown of National Merit Semifinalists. ❑

UNIVERSITY TESTING

Misused Hodgepodge of Tests Will Determine Students' Futures

The U.S. Department of Education (DOE) has set arbitrary cut-off scores on an unrelated group of multiple-choice tests to attempt to determine students' "ability to benefit" from post-secondary education. Such cut-off scores generally are considered a test misuse, even according to test-makers' guidelines.

The DOE's list of accepted tests ranges from the SAT, a college admissions test, to the Armed Services Vocational Aptitude Battery (ASVAB), to the Differential Aptitude Battery, which is primarily used for career planning. On each approved test, the cut-off score is set at "one standard deviation below the mean." For some exams, such a cut-off score falls near the *average* for minority groups.

The imprecision of multiple-choice tests makes cut-off scores capricious and unfair. On the SAT, for example, the margin of error is 68 points. According to the National Commission on Testing and Public Policy, since disadvantaged groups, on average, score nearer to the cut score, they bear the brunt of test inaccuracy. The culturally-bound nature of multiple-choice tests, which partially accounts for the generally lower scores of these groups, provides further argument that they should not be used in this way.

A review of the list of selected tests reveals carelessness in the regulations under which the law would be carried out. For example, the list includes an edition of the ASVAB which is "obsolete," according to a review by Dr. Lawrence M. Rudner, Director of the federal Educational Research Information Center (ERIC) Clearinghouse on Tests, Measurement, and Evaluation. Rudner also claims that for the only non-English test, the Spanish Assessment of Basic Education (SABE), "The materials... are designed for younger students. The specific stories and passages used in the SABE may be too childish, and hence offensive, for older students." ☐

Ability to Benefit

(Continued from page 16)

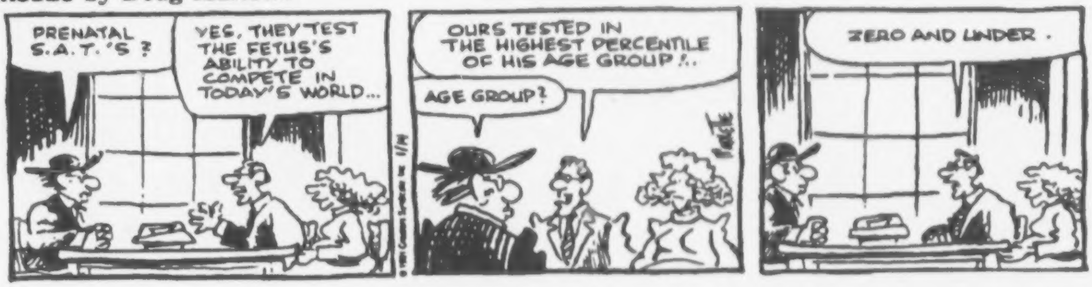
federal court in mid-January, and a technical amendment was passed by Congress in March to correct the error. The testing requirement still applies to students in need of financial aid.

The regulations also stated that institutions enrolling students who had not passed such a test would lose their federal financial aid.

In Washington, several bills have been proposed to address Ability to Benefit test misuses. Representative George Miller (D-CA) has drafted H.R. 907, that would allow diagnostic assessment and subsequent remedial education to meet the Ability to Benefit standard. Representative Ronald D. Coleman (D-TX) has proposed H.R. 533, which would repeal the provisions of the November law.

In March, FairTest and the California Community Colleges held an informational meeting on Ability to Benefit testing issues for education and civil rights groups in Washington, D.C. Many of these organizations expect to participate in Congressional hearings on the issue later this year. ☐

KUDZU by Doug Marlette


PRENATAL S.A.T.'S ?
YES, THEY TEST THE FETUS'S ABILITY TO COMPETE IN TODAY'S WORLD...
OURS TESTED IN THE HIGHEST PERCENTILE OF HIS AGE GROUP !..
AGE GROUP?
ZERO AND UNDER .

By permission of Doug Marlette and Creators Syndicate

UNIVERSITY TESTING

College Board's SAT Claims Discredited

"Validity Study Service" Reforms Urged

New research published in the May 1991 *Harvard Education Review* (*HER*) calls into question the College Board's widely-publicized "Dartmouth Study," which concluded that SAT scores predict first year college grades better than do high school grades.

According to *HER* authors James Crouse and Dale Trusheim (who co-authored *The Case Against the SAT*, University of Chicago Press, 1988), the Dartmouth study is "incomplete and slanted to promote the SAT..." Crouse and Trusheim claim that the Dartmouth analysis is flawed by its selective use of data, which "pertains only to differences observed between two groups: students admitted by high school record and rejected with the SAT, versus students admitted by the high school record and SAT and rejected with high school record alone. A complete interpretation of the data should include the students rejected by both policies and accepted by both policies." When such an investigation is done, according to the authors, it becomes clear that grades are still a better predictor of academic achievement and that the addition of the SAT to the high school record results in only a slight improvement in predicting average college grades.

Contrary to the College Board's claims, Crouse and Trusheim's re-analysis of Dartmouth admissions data found that adding the SAT to the high school record would cause virtually no change in the average freshman GPA, the number of academic disciplinary actions, or the number of academic honors given to students. Similar results were found for the more than 500 other colleges they evaluated.

The *HER* article focuses primarily on proposed improvements to the College Board's Validity Study Service (VSS), which assesses the selection benefits of the SAT. Crouse and Trusheim say VSS's methods "are significantly flawed, and this leads colleges to misleadingly positive conclusions" about the SAT. They suggest adding two tables to VSS reports to provide colleges with data showing "how much dropping the SAT and relying on only the high school record might hurt admissions decisions." They argue, "If a college that used the high school record and SAT were to admit and reject the same students with or without the SAT, dropping the test could not harm the college's admissions decisions."

The two suggested additions include a "Crosstabulation of Predicted Grades" and a table predicting "College Outcomes." The first would show the extent to which a college would admit the same students regardless of whether it uses the high school record or the high school record plus SAT scores. The second would show to what degree adding the SAT to the high school record improves the rate of college graduation, the average high school grades of admitted students, and the percent of admitted students with average first-year grades above 2.5.

While such information would be very revealing for most institutions, the authors say, "It is clear that the changes we recommend in the VSS will not be considered by the [College] Board or ETS unless colleges demand them. Colleges that use the VSS should inform ETS and the Board directly that these small changes to the VSS ought to be implemented."

Crouse and Trusheim conclude, "The important issue is not just whether redundant SAT information is important to admissions officers, or perhaps even to applicants themselves, but whether it is important enough to justify 1.7 million students taking the test each year." They argue that it is not. "As things stand now, it seems to us that the SAT is a monument to institutional inertia. Hundreds of colleges require applicants to take the test, but do not really use the test results."

▼ Copies of the *Harvard Educational Review* "Crouse article"(Vol. 61, No. 2) can be ordered by sending $3.50 to HER, 349 Gutman Library, 6 Appian Way, Cambridge, MA 02138. ❑

Subscribe to the *FairTest* Examiner

The only newsletter devoted entirely to testing reform issues

FairTest
EXAMINER

See page 15 for details

National Testing

Debate over national testing is intensifying this spring: President Bush proposed national testing, both houses of Congress held hearings, and dozens of groups and individuals have weighed in on both sides of the issue. If implemented, the national tests would be one of the most fundamental changes ever in the federal role in US education.

The articles that follow summarize Bush's plan, the Administration's efforts to avoid Congressional authority over national testing issues, the Congressional hearings, an interim report on the National Assessment of Educational Progress by the American Academy of Education, and the testing proposal of Educate America.

Essentially, two types of plans are proposed. One would be a predominantly multiple-choice test; the other would be a national examination "system" built around a set of national standards (see *Examiner*, Vol. 5, No. 1). President Bush's proposal incorporates elements of both plans. ❑

President Calls for National Test

As part of their *America 2000: An Education Strategy*, released April 18, President George Bush and new Education Secretary Lamar Alexander have called for national tests. Education Department budget revisions to fund the new exams were proposed in early May, and legislation, was filed in late May.

In response to the plan, FairTest Executive Director Cinthia Schuman said, "President Bush's proposal to rush into national testing deserves a flunking grade for putting politics over genuine educational reform." FairTest opposes federal legislation or appropriations for a national test or examination system at this time.

The plan calls for the development of "American Achievement Tests" based on "New World Standards." These tests would be developed "in conjunction with the National Education Goals Panel," created after the Charlottesville, VA education summit of September 1989.

Use of NAEP

Because *America 2000* authors recognize such tests "probably cannot be perfected" by 1994, they "will ask Congress to authorize the rapid deployment of an individual version of tests used by the existing National Assessment of Educational Progress." The idea of using NAEP below the state level has been widely opposed.

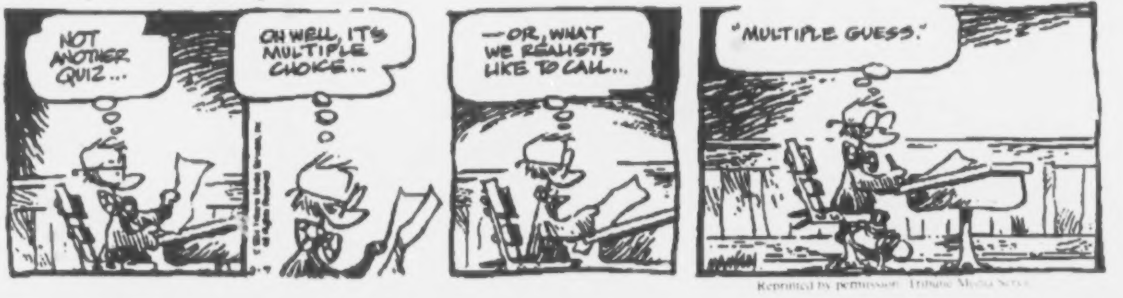
In a Goals Panel "Discussion Document" released on March 25, the Panel's Resource Group on Student Achievement recommended that NAEP be used for program monitoring and that "it be maintained in its current matrix sampling form." If used for individual testing, the matrix sampling method -- which involves testing a representative sample of students, each of whom takes only a small part of the test -- would have to be dropped.

America 2000 did not indicate what kinds of assessment an individualized NAEP might employ. The Campaign for Genuine Accountability "Statement on Proposals for a National Test" warned that "because of time and cost factors," rapid development of a national test will lead to a multiple-choice test, not use of performance-based assessments. In turn, this would "ultimately undermine public education."

Educate America

An individualized NAEP test could resemble the model proposed by Educate America, a group headed by former New Jersey Governor Tom Kean, which has called for establishment of a national test in six subject

SHOE by Jeff MacNelly


NOT ANOTHER QUIZ...
OH WELL, IT'S MULTIPLE CHOICE...
—OR, WHAT WE REALISTS LIKE TO CALL...
"MULTIPLE GUESS."

Reprinted by permission: Tribune Media Services

National Test Cont.

for every high school senior. Educate America claims these tests will be "state of the art," but will cost only $30 per student for all six. However, the SAT costs $16 for an entirely multiple-choice test. Since performance-based tests cost more to design and score, it is unlikely that the Educate America test could be anything other than almost entirely multiple-choice.

An Examination System?

The Goals Panel discussion document called for "Creating a National Educational Standards Framework" that would be used to develop a national examination system. This proposal is based on plans developed by the National Center for Education and the Economy (NCEE) and the Learning Research and Development Center (LRDC), whose director, Lauren Resnick, chairs the Goals Panel's Resource Group on Student Achievement.

Under the resource group's model, which the full panel has not yet voted on, "States or clusters of States would develop curriculum frameworks and examinations that would be calibrated to national anchor examinations based directly on nationally defined standards." As proposed by the NCEE/LRDC, this national exam system would rely on performance-based methods, not multiple-choice tests.

▼ *America 2000* is available free by calling 1-800-872-5327. "Measuring Progress toward the National Education Goals," a discussion document, is available free by calling (202) 632-0952. ❑

Avoiding Congress?

The Bush Administration plans to spend millions of dollars to initiate national testing, without Congressional authorization or public debate, even though imposition of a national test would be one of the most important changes in the history of federal involvement in education.

For activities wholly or largely related to new tests in the coming fiscal year, Education Secretary Lamar Alexander is requesting a total of $41.5 million more than Bush's original budget request. The funding will come largely at the expense of higher education programs.

In May, Alexander sent to Rep. William Natcher (D-KY), Chair of the House Appropriations Subcommittee on Education, a revision of the Administration's budget proposals for Fiscal Year 1992 (FY92, which starts Oct. 1, 1991). In it, he said, "Work on standards and tests is beginning under current authority and funding" for both the American Achievement Tests (the first of which is planned for 1993) and the modified NAEP.

FY91 funding is being used to construct the tests despite the fact that no authorization for any such national test exists, except for NAEP where legislation prohibits releasing individual scores or making comparisons below the state level. In some cases, funding earmarked by Congress for other purposes has been diverted without Congressional authorization.

The Administration's budget planners also claim that no additional authorization is needed for FY92 to launch its testing program, only appropriations under existing accounts. By this strategy, the Administration hopes to avoid the Senate and House committees governing education policy. Several key members of these bodies have already indicated opposition to a national test.

These tests would be carried out under "competitive grants or contracts," which would largely remove them from Congressional and public scrutiny. This could be done with the other exams as well, such as the American Achievement Tests. ❑

Quotes of the Month

"It seems to me you don't fatten the cattle by weighing (testing) them. You've got to feed them. And there's a lot of weighing going on here." — *Michael Casserly, Council of Great City Schools.*

"I continue to support such a system [of national examinations]... but it could also be a disaster. Will there be a mad rush to get exams in place whether they're good or not? And will they just be more of these standardized, multiple-choice, machine-scorable examinations? If so, we'll give a federal imprimatur to what is a very bad practice."— *Albert Shanker, American Federation of Teachers.* ❑

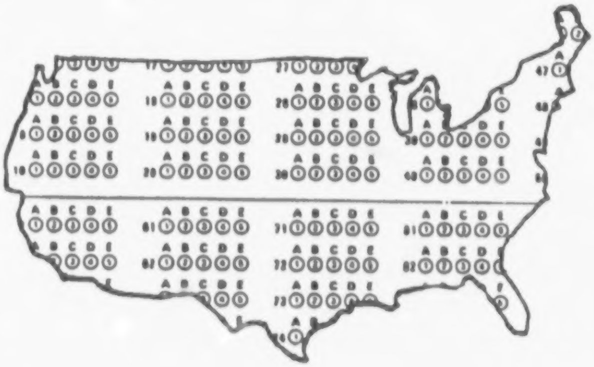
Congress Holds Hearings on National Test

Reacting to national testing proposals, both the US Senate and House have recently held hearings. FairTest Associate Director Monty Neill gave invited testimony at three of these sessions.

On March 7, the Senate Subcommittee on Education heard witnesses debate proposals for performance-based national examinations. Proponents included Lauren Resnick of the Learning Research and Development Center, Marc Tucker of the National Center on Education and the Economy, Gordon Ambach of the Council of Chief State School Officers, William Kolberg of the National Alliance of Business, and Albert Shanker of the American Federation of Teachers.

Critics, including the National Education Association, the American Association of School Administrators, the Council for Educational Development and Research (CEDaR), FairTest, and the Association of American Publishers, opposed or urged caution in developing a national asessment system, and strongly opposed a single test and multiple-choice testing. Proponents of the examination system also criticized multiple-choice testing.

In response to the idea of a performance-based national examination system, FairTest recommended that the Senate support the development of standards and implementation of performance-based assessments as part of a wider school reform effort focused on states and districts. But FairTest cautioned against adoption of a national exam system until more is known about the use of large-scale performance-based assessments, standards have been widely agreed upon, and questions of equity, governance, and educating teachers to use the new assessment methods have been adequately addressed. Even some proponents of a national exam system agreed that Congress should not legislate such a system.



House Hearings

The discussion before the House Subcommittee on Elementary and Secondary Education on March 14 was similar to the Senate hearing. Colorado Gov. Roy Romer, Chair of the Goals Panel, and William Brock, head of the Department of Labor Secretary's Commission on Achieving Needed Skills (SCANS) supported a national exam system. Gregory Anrig of ETS spoke in opposition to a national test, but indicated support for the ideas behind a national examination system.

On April 22, the House Subcommittee on Select Education also heard testimony on national testing, this time with a focus on the potential impact on minorities. Among those testifying, in addition to FairTest, were Eva Baker from the Center for Research on Evaluation, Standards and Student Testing; Joan Baratz-Snowden from the National Board for Professional Teaching Standards; Don Barfield from CEDaR; and Edward De Avila from Linguametrics.

Several witnesses pointed out potential dangers to students from low-income and minority-group backgrounds. They urged the federal government to assist in the development of performance-based assessment as part of supporting school reform, but cautioned against mandating an examination system at this time.

In his testimony, George Madaus, Boston College professor and director of the National Commission on Testing and Public Policy, criticized the use of European models to support US national test efforts as "misleading." He concluded with a strong metaphor: If we were discussing a medical treatment, we would first consider the consequences; we would perform careful trials and then weigh the pros and cons. Testing is an educational treatment, and we should do no less for the education of our children.

▼ For a copy of FairTest's written testimony to the House Select subcommittee, send a $.75 SASE to House Testimony, FairTest. Madaus' full report will be released by the Congressional Office of Technology Assessment this fall. ❑

279

The Future of NAEP

On March 13, the House Subcommittee on Elementary and Secondary Education held a hearing on the National Assessment of Educational Progress (NAEP). NAEP is a federal, biannual sample of student progress in various subject areas at three grade levels. In 1988, NAEP test-score comparisons among states were authorized on a trial basis.

At the hearing, Daniel Koretz of the RAND Corporation and Robert Linn, from the University of Colorado and co-chair of the study on NAEP state comparisons, strongly emphasized that NAEP should not be used to monitor individual student achievement. Doing so, they explained, would only lead to teaching to NAEP tests, inviting the "Lake Wobegon" effect of inflated test scores that do not correspond to actual student learning.

Opposition to obtaining individual NAEP scores was repeated in the "April 1991 Interim Report on the Evaluation of the 1990 NAEP Trial State Assessment," conducted by a National Academy of Education panel, as required under the 1988 legislation. The panel also recommended against extending NAEP to the school or district level without careful consideration of the consequences and costs. They expressed concern that use of NAEP at the school or small-district level "could lead to the loss of NAEP as an independent and uncorrupted indicator of educational progress."

The Academy panel did recommend release of the "state-level 1990 NAEP mathematics scale scores as scheduled," in June 1991. However, "because of serious concerns about the validity of the achievement levels developed last fall by NAGB [National Assessment Governing Board] the panel recommends... an independent technical review... prior to adoption, use or reporting of achievement levels."

The panel endorsed continuing the trial state assessments in 1994. Because "substantial lead time is required for achieving national consensus on new content frameworks, and for developing assessment questions and exercises that elicit more than rote learning from students," the panel recommended that authority for continuation "be made at the earliest time and that Congressional appropriations be at a level that will support appropriate assessment innovations."

The Panel also issued cautions on interpreting NAEP data, urging that conclusions be limited to those the data could support and that wider inferences not be drawn hastily. ❑

CAMPAIGN FOR GENUINE ACCOUNTABILITY

Arnold Fege of the National PTA answers questions from the media at the Campaign for Genuine Accountability March 27 press conference.

BEST COPY AVAILABLE

Myths of International Testing

Proponents of a national test for American schoolchildren often argue that the US needs a national test to set standards and stay economically competitive with Japan, Germany and other European countries.

However, according to testimony presented by Cornell professor E.W. Kelley to the House Subcommittee on Elementary and Secondary Education, none of these nations has a national test or examination system of the sort proposed for the US. Therefore, if these nations have superior educational systems, it cannot be because they have national testing.

Kelley explained, "Germany has no national assessment or any national test that almost all students take... Like the United States, control over educational instruction and evaluation lies principally at the state... and local level."

Furthermore, "No Japanese exam is designed to assess their education of all students or give the country a report card." (The Association for Supervision and Curriculum Development reports in the March *ASCD Update* that 93% of Japanese students do take an upper secondary entrance exam, and a third take a college entrance test.)

In France, Kelley continued, the baccalaureate "is still taken by fewer than 50% of students in their late teens." The baccalaureate consists of a number of different courses and subject-area exams.

Referring to these nations and Britain, Kelley observed, "None of these countries conduct internal assessments like our NAEP... None of these countries has a single test that all students take."

Kelley noted in his analysis of the Japanese system that a heavy emphasis on factual information and rote skills results from constant use of multiple-choice standardized exams to assess students, while "those more complex problem solving abilities that are most valuable to both student and potential employer" are not measured.

Kelley also discussed the commonly reported claim that the US lags behind other nations in test scores. He found that when actual percentages of test-takers are adjusted (the US top 10% versus other nation's top 10%), scores on subject tests are similar. He did suggest, however, that the lowest third of the US education system receives a particularly low-quality education. ❑

Exam Scam in Japan

A recent case of cheating by school officials has shocked and scandalized Japan. Under pressure from prominent alumni, school officials at Hyogo Prefectural Agricultural School altered several test booklets to ensure the success of several applicants to the 10th grade. But they botched the job and were caught immediately.

According to the Washington Post (3/30/91), "The shocking crime has gotten front-page banner headlines for days on end. Magazines are competing furiously to dig out every lurid detail. The now familiar videotape of the culprits being led to jail in handcuffs is shown virtually every night on the TV news." ❑

Groups Oppose Nat'l Test

(Continued from page 1)

Ann Lynch, National PTA President, added, "There are alternative measurements that will provide a more meaningful, more accurate assessment of student progress without creating a new national test."

The Statement urged Congress and the Administration to assist states and districts in implementing performance-based methods of assessment and improving staff development, as well as to support subject area groups in developing and disseminating model curricula, standards and assessments.

Sue Bredekamp of the National Association for the Education of Young Children explained, "The establishment of a national test before the curriculum debates have taken place will mean that the test itself will determine a national curriculum."

The Statement also urged that the National Assessment of Educational Progress remain a national monitoring system; not be used to compare students, schools or districts; and become increasingly performance-based.

Other groups endorsing the "Statement on Proposals for a National Test" include: American Association of School Administrators, ASPIRA Association, Center for Women Policy Studies, Mexican American Legal Defense and Educational Fund, National Association of Secondary School Principals, National Coalition of Title I Chapter I Parents, National Education Association, and Puerto Rican Legal Defense and Educational Fund. (For a complete list of signers, write to FairTest). ☐

Write to Congress

Now is the time to let Senators and Representatives know your thoughts about national testing.

FairTest recommends you address four points in your letter:

▼ Urge Congress to prohibit the Administration from initiating a national test or examination system using any federal funds not authorized for testing.

▼ Oppose legislation authorizing or appropriating funds for the creation of a national test or examination system at this time.

▼ Oppose creation of individualized tests and scores based on National Assessment of Educational Progress (NAEP) exams or items.

▼ Endorse Congressional support for the increased use of performance-based assessments as part of state and district school reform efforts.

✎ Send letters to both your Senators, any members of the House Education and Appropriation Committees who are from your state, and your own Representative. Write to Representative _____, Washington, DC 20515; or Senator _____, Washington, DC 20510.

U S House of Representatives Committee on Education and Labor

William D. Ford (MI), Chair
Joseph M. Gaydos (PA)
William Clay (MS)
George Miller (CA)
Austin J. Murphy (PA)
Dale E. Kildee (MI)
Pat Williams (MT)
Matthew G. Martinez (CA)
Major R. Owens (NY)
Charles A. Hayes (IL)
Carl C. Perkins (KY)
Thomas C. Sawyer (OH)
Donald M. Payne (NJ)
Nita M. Lowey (NY)
Jolene Unsoeld (WA)
Craig A. Washington (TX)
Jose Serrano (NY)
Patsy Mink (HI)
Robert A. Andrews (NJ)
William J. Jefferson (LA)
John F. Reed (RI)
Tim Roemer (IA)
Peter J. Visclosky (IA)
Ron De Lugo (VI)
Jaime B. Fuster (PR)
William F. Goodling (PA)
E. Thomas Coleman (MO)
Thomas E. Petri (WI)
Marge Roukema (NJ)
Steve Gunderson (WI)
Steve Bartlett (TX)
Richard K. Armey (TX)
Harris W. Fawell (IL)
Paul B. Henry (MI)
Cass Ballenger (NC)
Susan Molinari (NY)
Bill Barrett (NE)
John A. Boehner (OH)
Scott L. Klug (WI)

U S Senate Committee on Labor and Human Resources

Edward M. Kennedy (MA), Chair
Claiborne Pell (RI)
Howard M. Metzenbaum (OH)
Christopher J. Dodd (CT)
Paul Simon (IL)
Tom Harkin (IA)
Brock Adams (WA)
Barbara A. Mikulski (MD)
Jeff Bingaman (NM)
Paul Wellstone (MN)
Orin G. Hatch (UT)
Nancy Landon Kassebaum (KS)

(Continued on page 14)

Texas Weakens Retention Policy

In a partial reversal of its 1990 policy barring retention in pre-kindergarten and kindergarten, the Texas Board of Education has adopted a rule allowing students to be kept back one time prior to fourth grade, provided it is done with parental consent.

The new ruling also allows children to be placed in transitional programs with parental consent if a school deems this necessary for children to be "developmentally ready" for first grade. The Board still views transitional class placement as a form of retention.

One danger of the new ruling is that parents may be advised to allow their children to be retained or placed in transitional classes based on erroneous scores from standardized tests. Research has shown that retention of young children usually is more harmful than helpful.

A study of transitional classes in Texas by Mary Martin Patton revealed that in districts where these classes exist, the retention rates are 8 percent higher than those without them, and a disproportionate majority of those retained are males and minorities (*Education Week*, April 10, 1991).

Retention Findings Show Harmful Effects

Findings of a two-year study of children in transitional programs in Uinta County, Wyoming, show little achievement gain for students who spend an extra year in a transitional program. According to the study, differences were insignificant between "developmentally slow" children placed in transition classes before first grade and similarly labeled students who went directly to grade one.

Students in transitional programs were compared with a control group in language, math, reading, social skills, self-esteem, and rate of placement in special education. Those who had spent an extra year in school did not perform significantly better despite being a year older on average than children in the control group. However, the report did show significant difference in aggression between the two groups, with the transitional group being rated as more aggressive than the control group.

As with other studies of its kind, the author, Philip Ferguson, concludes that the differences in rates of development among young children even out over time, usually by grade 3 or 4 (*Education Week*, April 17, 1991). ☐

Appropriate Curriculum and Assessment for Young Children

The National Association for the Education of Young Children and the National Association of Early Childhood Specialists in State Departments of Education have issued "Guidelines for Appropriate Curriculum Content and Assessment in Programs Serving Children Ages 3 Through 8."

Two years in the making, the document is designed to "guide teachers and supervisors to: (1) make informed decisions about appropriate curriculum content and assessment, (2) evaluate existing curriculum and assessment practices, and (3) advocate for more appropriate approaches."

There are 18 assessment guidelines the primary purpose of which is "for planning instruction and communicating with parents." Some examples are:

- Assessment is relevant to, and a regular part of the curriculum;
- Assessment results are used to benefit the child;
- Each child is observed over time and in different circumstances so that assessment is representative of behavior;
- Assessment is not limited to one form but uses a variety of methods that may include artwork, written stories, lists or tape recordings of books read, systematic records of observations, conversations and interviews with children, and teachers' summaries of children's progress.
- Children's experience using their native language, their stage of development in the acquisition of language, and their ability to use English are all taken into consideration.

▼ The guidelines appeared in the March 1991 issue of Young Children and are available from NAEYC, 1834 Connecticut Ave. NW, Washington, DC 20009, (202)232-8777. ☐

K-12 TESTING

New York Campaign

The New York Campaign for Genuine Accountability in Education has published a Spanish version of its testing reform guide, *Standardized Tests and Our Children: A Guide to Testing Reform in New York*. With the release of *Los Exámenes Estandarizados y Nuestros Niños: Guía Para la Reforma de Exámenes Estandarizados en Nueva York*, more residents of New York will have access to easy-to-read information about standardized tests. The Campaign is a FairTest-initiated coalition of individuals and organizations working together to stop the inappropriate testing of children in the primary grades.

Following the October 1990 release of the Campaign's "Proposal to Replace Standardized Tests with Appropriate Assessment in the Early Childhood Grades," a series of information-sharing meetings took place between Campaign representatives and members of the State Education Department (SED) and the Board of Regents. Other recent activities of the Campaign have included regional events to celebrate the "Week of the Young Child," sponsored by the New York Association for the Education of Young Children (AEYC), and additional presentations to parent groups and educational organizations.

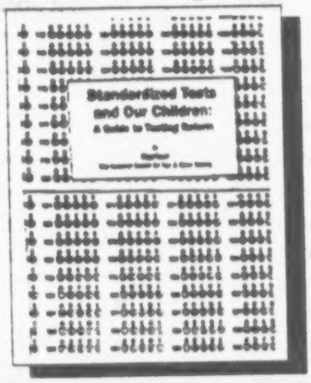
In New York City, the Campaign has also been working with community school boards to persuade them to support a ban on mass standardized testing of children from pre-kindergarten through grade 2. Staten Island's District 31 community school board recently joined District 15 and others endorsing the Campaign's "Statement to Urge a Ban on Standardized Testing of Young Children."

Because budget cuts are having a devastating effect on public education in New York City, testing has become a fiscal issue, as the wisdom of using limited funds to support unnecessary testing is called into question by educators, parents and community organizations.

▼ *Los Exámenes Estandarizados y Nuestros Niños (NY)* is available for $3 from the New York Campaign for Genuine Accountability, c/o Joan Harris, NYPIRG, 9 Murray Street, New NY 10007, (212) 349-6460. Bulk purchase discounts are available. ❑

Standardized Tests and Our Children:

A Guide to Testing Reform



Now available in Spanish and English

See page 15 to order

FairTest Training at SACUS

FairTest Field Organizer Veda Wright conducted a training workshop on advocating for developmentally appropriate assessment at the 42nd Annual Conference of the Southern Association for Children Under Six (SACUS) in Atlanta, March 11-16. The purpose of the workshop was to help train advocates who could lead assessment reform efforts in their respective states. Representatives from ten SACUS affiliates participated.

Participants had read *Standardized Tests and Our Children: A Guide to Testing Reform* prior to attending the workshop, which provided a common base of knowledge from which to begin. The workshop itself included an overview of developmentally appropriate assessment; discussion by participants of what is happening in their states in these areas; and steps to take in advocacy work, e.g. deciding on the focus, forming coalitions and networks, and identifying allies, resources and obstacles. Participants also brainstormed to help develop a plan of advocacy based on actual circumstances. ❑

U S Senate Committee on Labor and Human Resources

(continued from page 11)

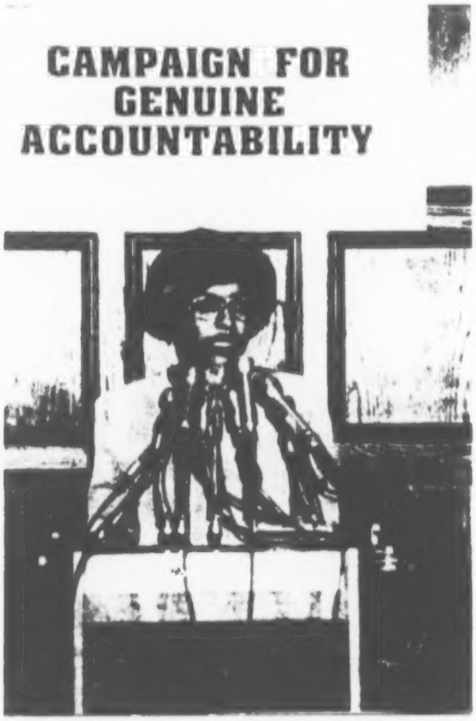
James Jeffords (VT)
Dan Coats (IN)
Strom Thurmond (SC)
Dave Durenberger (MN)
Thad Cochran (MS)

U S House Labor, Health and Human Services, and Education Appropriations Subcommittee

William H. Natcher (KY), Chair
Neal Smith (IA)
David R. Obey (WI)
Edward R. Roybal (CA)
Louis Stokes, (OH)
Joseph D. Early (MA)
Steny H. Hoyer (MD)
Robert J. Mrazek (NY)
Jamie L. Whitten (MI)
Carl D. Pursell (MI)
John Edward Porter (IL)
C.W. Bill Young (FL)
Vin Weber (MN)
Joseph M. McDade (PA), ex. officio

U S Senate Labor, Health and Human Services, Education Appropriations Subcommittee

Tom Harkin (IA)
Robert C. Byrd (WV)
Ernest F. Hollings (SC)
Quentin N. Burdick (ND)
Daniel K. Inouye (HI)
Dale Bumpers (AR)
Harry Reid (NV)
Brock Adams (WA)
Robert C. Byrd (WV), ex. officio
Arlen Specter (PA)
Mark O. Hatfield (OR)
Ted Stevens (AK)
Warren Rudman (NH)
Thad Cochran (MS)
Phil Gramm (TX)
Slade Gorton (WA)
Mark O. Hatfield (OR), ex. officio



Beverly Cole, Education Director, NAACP, speaks at March 27 press conference of the Campaign for Genuine Accountability.

Truth-in-Testing

(Continued from page 1)

FairTest played an active role in the case, *American Association of Medical Colleges v. Cuomo*, by filing a friend of the court brief in support of Truth-in-Testing. Ten civil rights, women's and education organizations also signed the FairTest brief, which argued that secret exams increase the possibility of biased questions and scoring errors, and leave test-makers completely unaccountable to the public. Truth-in-Testing was passed in New York in 1980 and extended nationally because of evidence that standardized tests often contain questions that are ambiguous and unfair.

The American Association of Medical Colleges (AAMC), makers of the Medical College Admissions Test (MCAT) and the plaintiff in the ten-year-old suit against the State of New York, have already attempted to overturn this testing reform victory. But in April, the Second Circuit rejected the test-maker's appeal for a rehearing and a hearing before the full Court of Appeals. Thus, unless AAMC appeals to the U.S. Supreme Court, the case will finally go to trial on the merits.

This victory for fair and open testing affects all university admissions tests, not just the MCAT. Following the District Court decision, sponsors of the Scholastic Aptitude Test, Graduate Record Exam, Graduate Management Admission Test and Test of English as a Foreign Language sought to stop New York from requiring that test questions and answers, along with documention that tests are fair and valid, be made available to the public. These test-makers have now agreed to partially comply with the law, pending the final outcome of the MCAT case. ❑

FairTest Publications

FairTest Examiner, published quarterly...$15.00/year (individuals), $25.00/year (institutions) $________

Back issues of *FairTest Examiner*, $4.00 each:
Vol____ No ____(All 14 for $40.00) $________

Standing Up to The SAT, by John Weiss, Barbara Beckwith and Bob Schaeffer (192 pp.)...$6.95 $________

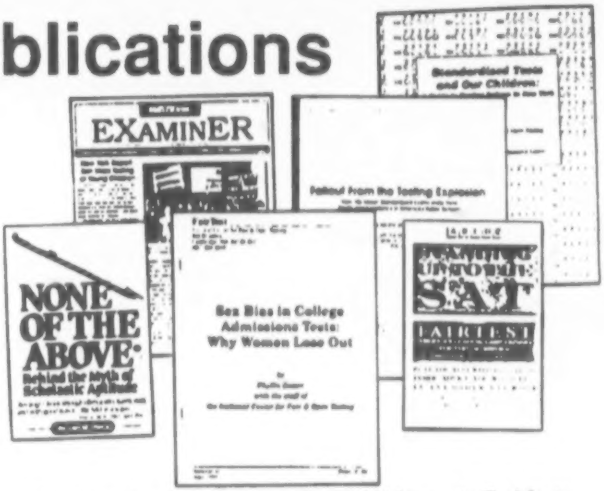
Fallout From the Testing Explosion: How 100 Million Standardized Exams Undermine Equity and Excellence in America's Public Schools, by Noe Medina and Monty Neill (77 pp.)...$8.95 $________

Sex Bias in College Admissions Tests: Why Women Lose Out, by Phyllis Rosser and FairTest staff, 3rd ed. (60 pp.)...$7.95 $________

The Reign of ETS, by Allan Nairn and Ralph Nader (550 pp.)...$30.00 $________

Beyond Standardized Tests: Admissions Alternatives That Work, by Amy Allina with FairTest staff (18 pp.)...$5.50 $________

None of the Above, Behind the Myth of Scholastic Aptitude, by David Owen (327 pp.)...$7.95 $________


EXAMINER
NONE OF THE ABOVE
Behind the Myth of Scholastic Aptitude
Sex Bias in College Admissions Tests: Why Women Lose Out
Fallout from the Testing Explosion
Standardized Tests and Our Children
SAT
FAIRTEST

Standardized Tests and Our Children: A Guide to Testing Reform, by FairTest (32 pp.)... $3; 5 for $10; 10 for $15; 50 for $50; 100 for $75. English and Spanish editions can be combined for bulk prices.

copies in English #________
copies in Spanish #________
Total $________

Standardized Tests and Our Children: A Guide to Testing Reform in New York, by FairTest & NYPIRC (32 pp.)...same prices as above. *Single* copies from FairTest. Bulk copies available from: NYPIRC, 184 Washington Ave., Albany, NY 12210

English #________
Spanish #________
Total $________

Contributions to FairTest

I agree there's no better way to be part of the testing reform movement than supporting FairTest! Enclosed is my tax-deductible contribution!*

☐ **FairTest Associate** ($25.00--includes a free subscription to the *FairTest Examiner*)

☐ **FairTest Sponsor** ($50.00--includes *Examiner* subscription and a FairTest publication of your choice, excluding *Reign of ETS*)

☐ **FairTest Sustainer** ($100.00 or more--includes above, plus a 20% discount on publications)

☐ **Other** ______

Publications Total $________
Contributions Total $________
Check Total $________

Name ______________________________
Address ______________________________
City ____________ State ______ Zip ________

Please make checks payable to FairTest and mail to: FairTest, 342 Broadway, Cambridge, MA 02139

*Contributions to FairTest are fully tax deductible, except for the value of publications received.

"Ability to Benefit" Testing Wipes Out Disadvantaged Students

New federal regulations that require students who do not have a high school diploma or General Education Diploma (GED) to demonstrate their "ability to benefit" from post-secondary education are barring disproportionate numbers of low-income and minority students from going to college. These students must now "pass" a standardized test before receiving federal financial aid.

At colleges as well as trade schools around the country, applicants are being turned away because of the new rules. "From years of experience, we know who will succeed here and who won't," said Dr. Dennis Brown, Communications Division Chair of El Paso Community College (EPCC) in Texas. "Many applicants are highly motivated to improve their lives, and many of them are the first in their family to go to college. To see their faces when they are turned away just because of their test score is devastating." At EPCC, about 240 students of the 2000 tested this spring were denied the opportunity to continue their education on the basis of a single test score. EPCC's population is about 80% Hispanic.

Other institutions, such as Bunker Hill Community College (BHCC), in Boston, Massachusetts, are finding their ability to serve communities of recent Asian immigrants hamstrung by the regulations. BHCC President Piedad Robertson told the *Boston Globe*, "It's just unbelievable. To me it translates to a conscious effort...to shut the door of financial aid on those who are most needy." BHCC's student population includes recent immigrants from Cambodia and Vietnam who do not have school documents proving they graduated from high school, so they must now be tested before enrolling. "They come here to learn English, but the first thing they have to do is pass a test that shows they know how to speak English." Robertson said.

Students object to Ability to Benefit testing as well. Testifying at a March Congressional hearing, Glen Becerra, President of the California Student Association of Community Colleges and a first generation college student from a farmworking family, said "We... fear that 'ability to benefit' students will have difficulty in accessing higher education if a test stands between them and federal financial aid....Students feel that it is wrong to use tests that have been proven to be culturally biased to determine a student's future."

The current controversy dates from last November, when the Ability to Benefit testing legislation was enacted. The law stated that, effective January 1, 1991, students without a high school diploma or GED must pass a U.S. Department of Education-approved standardized test in order to be eligible for federal financial aid. In December, shortly before the law went into effect, regulations were issued by the Department of Education stating that *all* students without a diploma or GED must demonstrate the "ability to benefit," not just students receiving federal financial aid. The faulty regulations were struck down by a

(Continued on page 4)

FairTest®
National Center for Fair & Open Testing
342 Broadway
Cambridge, MA 02139-1802
(617) 864-4810
(617) 497-2224 (FAX)

ADDRESS CORRECTION REQUESTED

Non-Profit
U.S. Postage
PAID
Cambridge, MA
Permit No. 53278

***Do you have an asterisk on your label?** If you do, this is your last issue, so **renew now!**

Moving? Return your label with new address

FairTest
EXAMINER
The National Center for Fair & Open Testing Vol. 5, No. 3 Summer, 1991 ISSN # 0898-2511

CAMPAIGN FOR GENUINE ACCOUNTABILITY

Steven Carbo of MALDEF, FairTest's Monty Neill, and Carol Choye of AASA at National Press Club news conference.

Campaign Opposes Bush Testing Plans

A FairTest-organized coalition of major national education, civil rights and advocacy groups has released an "Open Letter" calling on Congress to reject President Bush's plan for national testing of all U.S. schoolchildren.

At a June news conference in Washington, D.C. representatives of the Campaign for Genuine Accountability in Education urged Congress:

- Do not authorize or appropriate funds for development of a national test or examination system at this time;
- Do not legislate the development of a national examination system;
- Reject any plan to use the National Assessment of Education Progress (NAEP) to create individualized tests or provide individual student-by-student scores; and
- Support efforts of districts and states to develop and implement authentic, performance-based methods of assessment as part of wider efforts to reform curriculum and instruction and make a high quality education available to *all* our nation's children.

"A national test would be one of the most fundamental changes ever in the federal role in education," said Dr. Carol Choye, Superintendent of the Princeton (NJ) Regional Schools. Speaking as the representative of the American Association of School Administrators, Dr. Choye added, "Such an important policy decision should not be made without full, open debate by Congress and the public."

(Continued on page 7)

TM 017 189

Civil Rights Groups Ask Congress to Overhaul "Ability-to-Benefit" Rules

Leaders of major civil rights, women's and student groups have called on Congress to overhaul legislation that denies federal college aid to students without high school diplomas unless they "pass" a test proving their "ability to benefit" from post-secondary education. The ability-to-benefit testing law, enacted last November and amended this Spring, is denying an education to thousands of poor students across the country despite the lack of evidence that the tests will reduce student loan defaults, the apparent goal of the proposal.

(Continued on page 3)

CONTENTS

New NCAA Proposal 4
Deemphasizing MCAT 4
Franklin & Marshall Down Grades SAT 5
Testing Council Created 7
Performance Assessment 8
NAEP Math Results 12
Civil Rights Act 16

New Washington Staff Position Available

FairTest is establishing a Washington, D.C. satellite office to complement its testing reform agenda in elementary and secondary schools, higher education and employment. The Washington D.C. Coordinator will help serve as a liaison between the Washington community and FairTest staff.

Responsibilities include:

Start-up and management of the Washington, D.C. satellite office; implementation in Washington of FairTest's public education campaigns on various testing reform issues; research, monitoring and analysis of federal policies involving testing; and networking and coalition-building with national civil rights, professional and education organizations.

Qualifications:

FairTest is looking for a skilled public policy advocate and self-starter, with experience at the national level. Candidates should have strong organizational, strategic, speaking and writing skills. Ability to travel to, and work cooperatively with, the Cambridge office is essential. Policy experience in civil rights or education is preferred. People of color are urged to apply.

Salary: Depending on experience, plus health insurance and three weeks vacation.

To Apply: Send resume to: Washington, D.C. Coordinator Position, FairTest, 342 Broadway, Cambridge, MA 02139.

Starting Date: As soon as mutually agreeable. ❑

Carty-Bennia Internships A Success

This summer, FairTest launched the Denise Carty-Bennia Internship Program to honor the former Chair of our Board of Directors, who died in September, 1990. Two interns, Caroline Corbin, a recent graduate of Radcliffe College, and Jeff Lowenstein, an alumnus of Stanford University and Northeastern University's Graduate School of Education, worked on projects addressing racial, ethnic, class and gender bias in university admissions testing.

FairTest thanks both interns for their excellent work. Thanks also to the many friends and colleagues of Denise Carty-Bennia who contributed generously to this effort and helped make the first year of the program such a success. ❑

FairTest SAT Book Earns Top Library Award

Standing Up to the SAT, FairTest's popular guide to the Scholastic Aptitude Test published by ARCO/Simon & Schuster, has been honored as a "1991 Book for the Teen Age" by the New York Public Library.

The annual list, compiled since 1929, recognizes the best books in the nation written for young people ages 12-18. The booklist is distributed free to teenagers through New York Public Library branches and sold to over 3,000 other libraries in the U.S. and around the world.

▼ ***Standing Up to the SAT*** is available from FairTest for $6.95 postpaid. Use the order form on page 15 to order your copies today. It's great reading for teenagers and adults alike. ❑

FairTest

National Center for
Fair & Open Testing
342 Broadway
Cambridge, MA 02139-1802
(617) 864-4810 (617) 497-2224 (FAX)
FairTest EXAMINER, Summer 1991

Staff:
Cinthia Schuman, Esq., Executive Director
Monty Neill, Ed.D, Associate Director
Sarah Stockwell, University Tests Director
Veda Wright, Field Organizer
Bob Schaeffer, Public Education Director
Jamie Souder, Publications/Office Mgr.
Laura Pierce, Administrative Assistant

Board of Directors:
Ron Ellis, Esq., NAACP Legal Defense and Educational Fund; **Dr. Pamela George**, North Carolina Central University; **Deborah Meier**, Central Park East Secondary School; **Dr. Susana Navarro**; **Vito Perrone**, Harvard Graduate School of Education; **William Robinson**, Esq., D.C. School of Law; **Bob Schaeffer**, Treasurer; **Steve Solomon**, Esq., Proskauer, Rose, Goetz & Mendelsohn; **Chuck Stone**, *University of North Carolina, Chapel Hill;* **Ron Vera**, Esq., Best, Best & Krieger; **John Weiss**, Public Policy Initiatives, President.

National Policy Panel:
Dr. Leonard Beckum, Duke University; **Dr. Beverly Cole**, NAACP; **Dr. Harold Dent**, Association of Black Psychologists; **Arnold Fege**, National PTA; **Dr. Howard Gardner**, Harvard University; **Dr. Asa Hilliard**, Georgia State University; **Dr. Lucinda Lee Katz**, University of Chicago Laboratory Schools; **U.S. Rep. Matthew Martinez** (D-CA); **Sara Melendez**, University of Bridgeport; **Aurelio Montemayor**, Intercultural Development Research Association; **Harry Pachon**, National Association of Latino Elected and Appointed Officials; **Alfredo de los Santos**, Maricopa Community Colleges; **Eleanor Smeal**, Fund for a Feminist Majority; **Theodore Shaw**, Esq., University of Michigan; **Dr. Gail Thomas**, Texas A & M University; **U.S. Rep. Ted Weiss** (D-NY).

Funding for FairTest's work comes from **the Golden Rule Corporation, the Ford Foundation, the Rockefeller Family Fund, the Philip M. Stern Family Fund, the Aaron Diamond Foundation, the Edward Hazen Foundation, the Lilly Endowment, the New York Community Trust, and many individual contributors.**

FairTest and Civil Rights Groups Ask Congress to Overhaul "Ability-to-Benefit" Rules

(Continued from page 1)

In a statement of principles on "Standardized Tests in Federal Post-Secondary Education Policy," initiated by FairTest and delivered to House and Senate Education Committee members in July, the groups said, "Multiple-choice tests are often seen as simple, objective and inexpensive ways to solve complex problems, such as high student loan default rates or the abuse of the federal student loan system by some unscrupulous institutions. But, like many seemingly 'easy' answers, arbitrarily imposing minimum test-score requirements -- or 'cut-off scores' -- creates more problems than it solves."

Many of the multiple-choice tests used to determine ability-to-benefit -- such as the SAT, ACT and Armed Service Vocational Aptitude Battery (ASVAB) -- have been shown to be biased, inaccurate, or irrelevant predictors of future behavior, especially for students whose first language is not English. A well-known example of test inaccuracy involves the ASVAB: in 1976 an error in setting the cut-off score on the ASVAB resulted in 300,000 recruits being admitted into the military who would normally have been rejected. These enlistees, as a group, ended up performing about as well as those who "passed" the test.

"Ability-to-benefit is a Catch 22," said Dr. Dennis E. Brown, Communications Division Chair at El Paso Community College, one of the endorsers of the statement. "A non-native English speaker needs language training to successfully pursue college studies, but the ability-to-benefit legislation bars these students from obtaining federal financial aid to take those courses." At El Paso Community College, whose enrollment is 80% Hispanic, roughly 240 students of the 2000 tested last spring were denied an opportunity to continue their education.

Other endorsers of the principles include the Mexican American Legal Defense and Educational Fund, National Association for the Advancement of Colored People (NAACP), National Council of La Raza, Puerto Rican Legal Defense and Educational Fund, Center for Women Policy Studies, Association for Women in Science, and United States Student Association.

Like many seemingly 'easy' answers, arbitrarily imposing minimum test-score requirements -- or 'cut-off scores' -- creates more problems than it solves.

The groups cited these problems in the federal testing rules:

- Test score cut-offs are arbitrary by nature;
- The culturally-bound nature of tests makes cut-off scores unfair to students from non-English-speaking backgrounds;
- Tests should never be used alone to make 'high stakes' decisions about individual students or programs since they are imprecise instruments;
- The federal government should not create de-facto admissions standards by enacting test score minimums.

The current controversy began last November when ability-to-benefit testing legislation was enacted. In December 1990, shortly before the law went into effect, the U.S. Department of Education issued regulations requiring *all* students without a diploma or GED to demonstrate their ability to benefit, not just students receiving federal financial aid. The faulty regulations were struck down by a federal court in mid-January, and a technical amendment was passed by Congress to correct the error. However, ability-to-benefit testing is still required for students seeking financial aid. New regulations are expected by the end of the summer.

Rep. George Miller (D - CA) has drafted legislation that would allow assessment instruments that are used diagnostically (followed by remedial education if needed) to meet the ability-to-benefit standard for two- and four-year colleges. Rep. Ronald Coleman (D. - TX) has sponsored a bill repealing the entire law. FairTest and others submitted testimony on these bills at U.S. House Subcommittee on Postsecondary Education hearings in June. Further Congressional action is expected later this year.

▼ For a copy of the principles, "Standardized Tests in Federal Post-Secondary Education Policy," Send a SASE to FairTest. ❑

New NCAA Proposal to "Toughen" Academic Standards Is Flawed

The President's Commission of the National Collegiate Athletic Association (NCAA) has proposed changes in its athletic eligibility rule, Proposition 48, which effectively increase the emphasis on standardized test scores in the guise of "toughening" academic requirements.

The proposal would require every first year athlete at a Division I school to have a minimum test score of 700 on the SAT (or 18 on the ACT) plus a 2.5 grade point average (GPA) in 13 core courses in high school. Currently, Prop. 48 requires that students have a 2.0 GPA in 11 core courses plus the minimum test score. Though the test score minimum is unchanged in the presidents' plan, students with higher scores will be eligible with lower GPAs. For example, scoring 900 on the SAT would drop the minimum acceptable GPA to 2.0.

Another NCAA committee had proposed that test scores also be indexed. For example, an athlete who scored 400 on the SAT would become eligible with a 3.25 grade point average. This proposal was designed to lessen the NCAA's dependence on test scores in response to criticism that they disproportionately exclude African American athletes.

But the NCAA's President's Commission elected to stick with its arbitrary cut-off score, called "a line in the sand" by David L. Warren, president of Ohio Wesleyan University. "To go below that would be politically and educationally a retreat," Warren said. The presidents ignored data proving that higher test scores do not necessarily indicate success in college and that the high school record is the best predictor of college performance. Instead they relied on their false impression that low-scorers, even with impressive high school records, would fail in college. "There was a general feeling that students who score below 700 really don't have a chance of being successful on our various campuses," said William Cunningham, president of the University of Texas at Austin. (*Chronicle of Higher Education*, 7/3/91).

This "general feeling" is not borne out by the evidence. Students with a strong high school record generally do well in college, whatever their test scores. Research shows that even students with a moderate high school G.P.A. can succeed in college, regardless of their test scores. Rather than toughening standards, the NCAA presidents sent the message that four years of high school work are less important than scores on a biased, coachable, imprecise test taken in three hours on a Saturday morning.

The presidents' proposals will be considered by NCAA members at their annual convention next January. ❑

Medical Schools With High Minority Enrollments Rely Less on MCAT Scores

Medical schools which enroll and graduate minority students at higher rates than average place less emphasis on test scores in admissions, according to a recent study by the Southern Regional Education Board (SREB). These institutions tend to aggressively recruit minority students and are more likely to use test scores to indicate which students need special help rather than to exclude them. They then provide such students with academic and financial support. As a result, they graduate minority students at virtually the same rate (89%) as their overall student population (90%).

While college admissions tests such as the MCAT are designed to predict student success in their first year in college, medical schools which graduate minority students at high rates "define success not as a lack of academic problems during medical school, but rather as the ultimate completion of the curriculum and attainment of licensure. Essentially, they apply the test of success to the end product, the graduate physician, rather than to performance during the educational process."

Institution-wide sensitivity to and support for minority students was key to recruiting and retention. According to the study, admissions committees in particular need to pay attention to turning rhetoric into practice: "Convincing members of admissions committees of the need to judge applicants by non-traditional measures, and giving them confidence in their ability to do so, is a serious obstacle to increasing minority admissions."

Universities which enroll a smaller than average proportion of minority students (less than 11.4 percent), graduate only 75 percent of them, compared to 90 percent of all students. These schools tend to weigh test scores more heavily in the admissions process.

Johns Hopkins Medical School, which doesn't require the MCAT (it accepts scores from any admissions test instead), is one of the schools enrolling and graduating minority students at a higher-than-average rate.

▼ For copies of the report, *Recruitment and Retention of Minority Medical Students in SREB States*, send $ 7.50 to SREB, 592 Tenth St. NW, Atlanta, GA 30318-5790. ❑

C. Fisher

Franklin & Marshall Drops the SAT for Top Students

This spring, Franklin & Marshall College in Lancaster, Pennsylvania, joined the more than 100 colleges around the country which are deemphasizing test scores in their admissions process. Starting with students who are applying for the fall of 1992, the SAT will be optional for applicants in the top ten percent of their high school class. These students will have the option of submitting two graded writing assignments instead of test scores.

Currently, over forty percent of F&M's applicants are in the top ten percent of their class. F&M is a "highly competitive" college which admits fewer than half of its applicants.

According to Franklin & Marshall's Director of Admissions, Peter Van Buskirk, the change in policy stems, in part, from the growing nationwide dissatisfaction with the role of the SAT in the college admissions process. "A lot of schools are beginning to discuss some sort of test optional policy -- it's a hot topic," said Van Buskirk. "Many admissions officers have expressed frustration with the rankings of colleges by scores and with the fact that students' first question about a school invariably focuses on test scores.... In conversations with students and parents, our new policy has met with uniform approval. They don't have to work under the ominous burden of SAT scores anymore."

"We weren't convinced that the SAT is a necessary predictor for high achievers," he continued. "We find our best, most thorough indicators of potential for academic success in the high school transcript... Unfortunately, a relatively modest SAT performance can introduce an unfair bias in cases where the student has achieved extremely well in the classroom. We want to remove that bias from decision-making."

> ***"A lot of schools are beginning to discuss some sort of test optional policy -- it's a hot topic,"***
>
> ***-Peter Van Buskirk***
>
> ***F & M Director of Admissions***

The change in policy was proposed by Franklin & Marshall's Faculty Committee on Admissions after conducting several internal studies which proved that the high school record is the best predictor of academic work at F&M, especially for high achieving students. F&M hopes the new policy will increase the volume of applications from such students, who may have been discouraged from applying in the past because of SAT scores below F&M's average. According to a committee report, "For these students, the requirement to submit such a score may have been posing an unwarranted barrier, a disincentive, although they might well perform satisfactorily here." The report cited the example of Bates College, which experienced a boost in applications after deemphasizing the SAT in 1984. The committee will decide in 1994 if the new option should be expanded to all applicants. ❑

National Assessment of Young Children

The National Education Goals Panel is considering development of a national school readiness assessment to help determine progress toward the first of the six national goals adopted by President Bush and the Governors. Goal 1 states, "By the year 2000, all children in America will start school ready to learn."

The Panel's Advisory Group on Goal 1 has recommended, as indicators of "readiness," collection of data about children's health and nutrition, participation in pre-school activities, and parental background information. They also called for development of a "National In-School Assessment Profile," to be used on a sampling basis with individual children "to produce valid and reliable measures of student readiness... [in] physical well-being, emotional maturity, social confidence, language richness, and general knowledge." A task force set up to consider what such an assessment might look like is due to report this summer.

Task force members have already questioned whether the resources will be available to do the assessment properly. Child advocacy groups from across the nation have expressed fear that the value of the information will be outweighed by potential misuses such as sorting children into special programs based on test results. The task force therefore is likely to recommend proceeding only with great caution. ❑

AERA Cautions on National Testing

"The American Educational Research Association (AERA) has concluded that a preponderance of research evidence suggests that America slow down, and think differently, about test-based educational accountability," reads a resolution adopted by the AERA Executive Council in early June. The AERA is a membership organization of thousands of university, government and other educational researchers.

The resolution was approved the same week as an AERA-sponsored invitational conference on the issue of national testing. A panel of some of the nation's most distinguished researchers presented evidence clearly indicating that the wave of state minimum competency tests of the past 20 years has harmed, not helped, education.

Panelists Linda Darling-Hammond, Richard Jaeger, George Madaus, Robert Stake, and Lorrie Shepard presented a range of critiques about testing programs and their effects. Each drew on research to strongly oppose national multiple-choice testing, such as use of the National Assessment of Educational Progress to create tests that could be used to report individual scores.

Proposals for a national examination system that would allow for each state to develop its own performance-based assessments that would then be calibrated to national standards were not as directly opposed, but the panelists were clearly skeptical that the benefits would outweigh the dangers. Panelists agreed the process of developing national standards and assessments must be careful, and more knowledge is required about effects before final decisions on a national system can be made.

Commentators Gerald Bracey and Milbrey McLaughlin concurred with the findings of the presenters.

▼ The papers are available from AERA, 1230 17th St. NW, Washington, DC 20036. ❑

Nat'l Conference on Black Children

The **National Conference on Educating Black Children** has joined the growing list of groups opposing a national test. At the sixth annual conference, over 400 parents, teachers, clergy, school administrators and others met in Maryland in early June to discuss the impact of assessment and school-based management on black students.

Conference speakers on assessment addressed the disproportionately negative effects of standardized testing on black students. U.S. Rep. Major Owens (D-NY), chair of the House Subcommittee on Select Education, warned of the dangers of a national test, particularly the achievement tests in President Bush's *America 2000* plan.

FairTest Field Organizer Veda Wright led a workshop session examining testing and assessment. ❑

Congress Creates Testing Council

Congress has passed, and President Bush signed, legislation creating a National Council on Education Standards and Testing. The new Council is to investigate the desirability and feasibility of developing national education standards or national exams and report its recommendations by the end of 1991.

The Council combines a Congressionally-desired body to investigate the desirability of standards and testing with an Interim Council created by the National Education Goals Panel to investigate how to develop national standards and testing. The membership of the Interim Council was transferred to the National Council. Thus, while the stated goal of the Council is to study whether standards and testing are desirable and feasible, the Council majority may ignore the issue of desirability and simply proceed with feasibility questions. Indeed, its chair, Gov. Roy Romer, clearly expects to proceed with planning the development of standards and testing.

In addition to the appointees continuing from the Goals Panel, there are four members of Congress (Sens. Bingaman and Hatch, Reps. Kildee and Goodling), plus an additional four persons, not yet named, for a total of 32 members.

At the first Council meeting, held June 24, the Congressional members indicated concern about the desirability of national testing. Sen. Bingaman questioned whether the resources would be committed to allow all children to do well on the tests. Sen. Hatch wondered if it would lead to a national curriculum and reminded the Council that the statute creating the Dept. of Education stated there would not be a national curriculum. Rep. Goodling, the ranking minority member of the House Education Committee, expressed doubt that testing would lead to improved education.

The panel heard from witnesses representing subject area groups such as math and science. They consistently indicated that their organizations are developing and implementing standards for curriculum, teaching and assessment, but the process is complex and requires time for reflection and grassroots participation. Gov. Romer kept insisting that the process could be greatly hastened, using analogies from production in businesses such as auto manufacturing. Panelists both questioned the validity of his metaphors and indicated that while some aspects could be hastened by having additional funds, the process was necessarily time-consuming.

The Council also heard from three assessment experts, two of whom urged caution in the development of national assessments due to lack of knowledge about large-scale performance-based assessments, the effects of such assessments, or how they might be used in ways harmful to low-income or minority-group children. ❑

Campaign Opposes Bush Testing Plans

(Continued from page 1)

"Politicians cannot simply mandate new tests and expect education to magically improve," explained FairTest's Monty Neill. "Other reforms, such as new curricular standards, improved instructional methods, and financial resources must be in place first. But the Bush Administration does not plan to devote the resources needed to bring education up to high quality standards for all. Instead, they simply plan to test all children in grades 4, 8 and 12."

Steven Carbo of the Mexican American Legal Defense and Educational Fund warned, "Given the low correlation between test scores and future performance, the testing plan will become a gatekeeper for Hispanics and other minorities, closing off access to educational and employment opportunity."

The Campaign supports the development and implementation of performance-based assessments, added Jennifer Tucker of the Center for Women Policy Studies. "However, it is early in the process and there is much to be learned about how effectively and fairly these tests will assess our diverse student population. We must ensure that subtle biases do not find their way into performance-based assessments."

Other organizations signing the "Open Letter" include the Center for Law and Education, National Alliance of Black School Educators, NAACP, National Association for the Education of Young Children, National Elementary and Secondary Principals' Associations, National Council of La Raza, National Council for the Social Studies, National PTA, Puerto Rican Legal Defense and Educational Fund, and the Southern Association on Children Under Six.

▼ For a copy of the Open Letter and a list of signers, send a legal-size SASE to FairTest. ❑

Performance Assessments Continue to Develop

Performance assessments -- also referred to as authentic, appropriate or direct assessments -- are spreading across the US. In the past year, New York and Oregon have decided to switch to predominantly performance assessments, several other states have introduced elements of performance assessments, and others are considering doing so.

Those states that had already committed to the new assessments continued to make progress (see *Examiner*, Summer 1990). Vermont, Arizona, Maryland and Connecticut each piloted their new assessments, all but Maryland on a partial, trial basis. California is considering legislation to greatly expand the role of the California Assessment Project and the use of performance assessments. North Carolina's state-mandated performance assessments for grades 1 and 2 are now firmly in place. Kentucky has detailed its plans and awarded a contract to develop the actual exams.

Consortia of states and districts are also continuing to work on various aspects of new assessments. These include: a Chief State School Officers program; a science assessment consortium led by Connecticut; and the Joint Standards Project, a private effort supported by foundations and headed by Lauren Resnick and Mark Tucker that is attempting to develop a national assessment system.

This issue examines a few states and the progress they have made over the past year. This survey of state developments will continue in the next issue.

Arizona

The Arizona Department of Education has begun to introduce performance assessment into classrooms across the state. This year, teachers in grades 3, 8 and 12 were given sample forms of the new Arizona State Assessment Project (ASAP) to try out in reading, writing and math classes. In the spring of 1992, a different version of the test will be administered under exam conditions to provide baseline data for the state.

Detailed critiques from teachers about the assessments, ranging from inadequate directions to the need for more thinking skills, will be used to revise the forms this summer. The revisions will be most extensive in math in an effort to bring the Arizona assessments more in line with the standards of the National Council of Teachers of Mathematics. Spanish-language forms are also being developed.

The assessments are produced by Riverside Publishing under contract with the Arizona Dept. of Education. Scoring is done by teachers who are being trained to grade the assessments. The state is also helping districts develop portfolios and other methods of assessing at the local level to meet state requirements that each district develop an assessment plan to complement the ASAP.

Despite conferences designed to enable teachers to use new assessments, insufficient staff development has been a major problem. Some districts lacked funds to hire substitute teachers to free up regular teachers to attend the conferences. Another problem is the assessments' heavy reliance on writing ability, even in the math tests.

Most teachers have responded positively, even while recognizing the problems, saying such things as, "Finally we are being assessed on what we teach." Larger districts have tended to move more slowly, Education Department personnel explained, due to larger bureaucracies and more investment in multiple-choice testing.

For the coming school year, the state has reduced the use of its conventional "achievement" tests, the Iowa Test of Basic Skills (ITBS) and the Tests of Achievement Proficiency (TAP). In the past, districts were required to administer the test to every student in grades 2 through 12. Now they are mandated to test only in grades 4, 7 and 11. Districts must use the reading comprehension and math problem-solving parts of the ITBS or TAP, but may ignore the rest. The state will not pay for ITBS or TAP testing in other grades. Reducing testing to three grades was seen as an economy move by a cash-starved legislature, which will reconsider use of the ITBS next year.

The ASAP exams are meant to encourage changes in curriculum and instruction. In each subject area and grade, a variety of ASAP forms will exist. Each teacher will have available all the forms for his or her grade, except for the one actually used for exams.

The state will also create similar forms for each of the two grades preceding the tested grades. The use of all these forms is voluntary. Again, insufficient staff development has been a problem.

Several versions of the same form will also be used to administer each year's test in each subject. Scores on the different versions of the form will not be psychometrically equivalent because the items will differ.

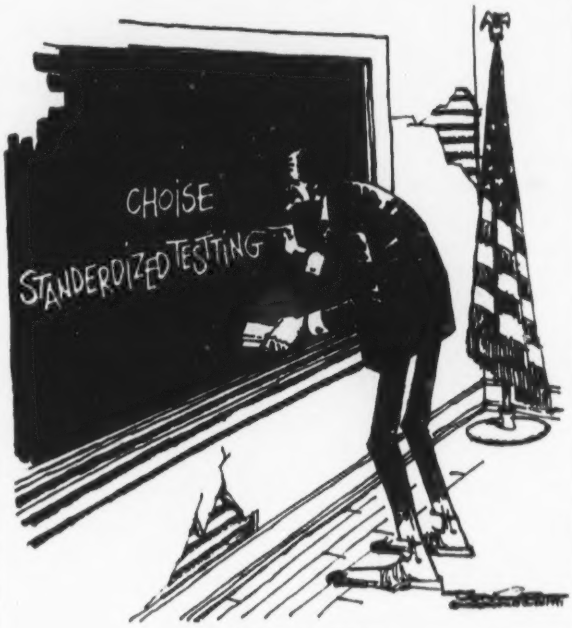
The issue of equivalence on performance exams has always been a difficult problem: if students work on different complex problems (each of which may take substantial time), how do we know each has demonstrated similar skill levels? This problem primarily exists if individual scores are reported or used for comparison or decision-making purposes. If they are designed to provide instructional information for individuals and aggregate data at the state level, the problem can be avoided.

Vermont

This past school year, Vermont piloted its new assessments in writing and math in grades 4 and 8. According to Ross Brewer, Director of Planning and Policy Development in the Vermont Department of Education, the biggest accomplishments were the development of standards to evaluate math portfolios and the identification of specific areas in which teachers need to be trained to be successful in implementing the new assessments.

One hundred and forty-eight schools participated in pilot assessments which included writing and math portfolios, student-selected best pieces in writing and math, and a uniform writing test in which students were given up to 90 minutes to respond to a writing prompt. No uniform math test has yet been developed.

Standards were developed entirely by teachers, though Advanced Systems, a testing company, has worked with the Department of Education to develop and refine the scoring process. A committee of teachers received training enabling them to score the assessments.



The Edukashun Prezedent

Teachers were generally more successful in using the writing assessment in the classroom than the mathematics, but they will begin receiving training from other teachers during the summer at 6-week long math institutes. Throughout the coming school year, regional training sessions run by teachers will also take place. Networks of approximately twenty teachers, each with a trained head teacher, will operate throughout the state to enable teachers to learn from each other and continue to develop standards.

The Vermont legislature has allotted $425,000 to implement the new assessments in the 1991-92 school year in the state's approximately 225 fourth and eighth grade classes, the same amount allotted for two-thirds as many classes in the pilot year.

Plans are being made for each school to hold a School Report Night when parents and other community members are invited to talk about student achievement, ask questions and look at students' work.

▾ For copies of the writing and math reports, contact Ms. Elizabeth Rand, Vermont Department of Education, 120 State Street, Montpelier, VT 05620.

North Carolina

North Carolina has completed the third year of its *Grades 1 & 2 Assessment in Communication Skills and Mathematics*. The Assessments rely on teacher observation and collection of student work, organized around a set of forms for indicating student progress.

Approximately 95% of the State's school systems have provided staff development for implementing the assessments. The North Carolina Department of Public Instruction offers regional training sessions where master teachers work with teachers from the districts, who then return to their home districts and work with other teachers in the use of the assessments.

During these trainings, teachers use real classrooms that are set up in developmentally appropriate ways. Among other things, they receive training to facilitate student learning through integrating various subjects into a theme, project or topic. They also receive help and support in learning how to make informed judgments about individual children,

(Continued on page 10)

(Continued from page 9)

an important step in the process of moving away from reliance on standardized tests. When there are problems, they are usually attributable to teachers not using developmentally appropriate practices.

For parents, "seeing is believing." Parents might be shown samples of work that are representative of children in second grade in the spring term, as well as their own child's work. When parents are able to see the growth in their child's work that occurs throughout the school year, they are convinced of the value of the assessments.

The record of each child that gets passed to the next year's teacher is a profile or portfolio containing the student's work and the major concepts or learning strategies he or she has mastered.

Three-day orientation sessions have been offered to people involved in higher education to acquaint them with the new assessments for which future teachers must be educated. Many have asked for further assistance and resources.

One remaining problem in implementing the assessments is the requirement to administer norm-referenced standardized tests in order to continue to receive Chapter I federal funds for disadvantaged students. The use of standardized tests with young children conflicts with developmentally appropriate curriculum and assessment and, thus, with the North Carolina approach.

Kentucky

Kentucky has awarded a contract to Advanced Systems to implement the Kentucky Instructional Results Information System (KIRIS). KIRIS was created in response to the Kentucky Education Reform Act of 1990, which called for a complete restructuring of that state's public education system, including the development of a "primarily performance-based assessment system."

KIRIS is scheduled to take effect in the spring of 1992 and is designed to direct instruction and assessment towards the achievement of six goals: basic skills, core concepts, self-sufficiency, group membership, problem solving and integration of knowledge.

The Accountability Assessment Component (AAC) of the system is mandatory and requires yearly assessment of all students in grades 4, 8 and 12 in the six goals. It will consist of tests similar to those administered by the National Assessment of Educational Progress (NAEP), which are mainly multiple-choice; performance assessments in which students must produce a product or solve a problem based on subjects taught in the classroom; and portfolios of each student's best work throughout the year.

The Continuous Assessment Component is voluntary and has the same elements as the AAC, but focuses on grades K-3, 5-7, and 9-11. It allows for assessment of daily classroom activities and is designed to assist instruction as well as prepare teachers and students for the AAC. The Accountability Assessments will be used to hold teachers, schools and districts accountable for ensuring continuous educational progress for their students.

The decision to utilize a NAEP-like test has been somewhat controversial because it will be predominantly multiple-choice. Moreover, NAEP items themselves are not available for use by any entity other than NAEP, and NAEP uses a sampling procedure that prohibits meaningful individual scores. The Bush Administration, however, is attempting to remove the prohibition on the use of individual NAEP items in order to develop individualized NAEP exams as part of its *America 2000* proposal.

Lessons from the States

The complexity and richness of the various efforts, as well as the many problems that remain to be solved, reinforce the view that a national examination system would be premature. Rather, developments in various states should be allowed to proceed, be studied, critiqued and revised. Each of these efforts, as well as those in California, Connecticut, Maryland and other states, districts and schools, can contribute to finding better ways to assess student learning and to use assessment results to improve education. ❑

NJ Resolution

The New Jersey School Boards Association adopted a resolution in June calling for a deemphasis on standardized tests and the development of multiple indicators of students' performance "including, but not limited to the evaluation of student portfolios, samples of students' actual work, teacher observations of students' classroom performance, as well as standardized test results." The Association plans to lobby the state commissioner of education and the legislature to develop an evaluation system that is not dependent on standardized tests. Also, a Task Force on Educational Assessment and Monitoring has been set up by Governor Florio to make recommendations on performance standards and assessment of those standards. ❑

Sparking Local Test Reform

Guest Column by Holly Y. Rechel-Felmlee

This spring, parents, teachers and administrators persuaded two rural school districts near Alamosa, Colorado to sharply reduce their standardized testing.

In Alamosa, about twenty families requested that the district not administer standardized tests to their children. Some schools in the district told parents to keep their children home during the testing time. Other schools, mostly primary, arranged activities at school for those children who would not be tested.

This act by local parents raised questions throughout the region about the need for standardized tests and the uses of test scores. Colorado only requires that districts test three times between kindergarten and twelfth grade, yet few parents knew this.

The public discussion began in the spring of 1990 when, as a local education columnist, I decided to attack the test issue head on. I wrote a column in the *Valley Courier* newspaper denouncing standardized tests, and said that I would be keeping my kindergarten child out of testing. I was told to keep my child at home. In order to keep her perfect attendance record, however, my child chose to attend school and was given the test.

This spring I considered dropping the test issue. I felt defeated by the bureaucracy of school and alone in my stand.

Then I was approached by several parents in the district (some of whom were also teachers), asking me to write again about testing. They said they would use my articles and their own research to talk to parents about standardized tests.

I wrote two columns against standardized tests and encouraged parents to write to their principals and superintendents specifically requesting that their children not be tested, if that was their choice.

Many teachers and administrators in Alamosa and surrounding districts, as well as some parents, were already concerned about standardized testing, its validity to the education process, and its use by their district.

Holly Y. Rechel-Felmlee

My columns reinforced the message from some administrators to their school boards that testing was an important issue and that many organizations had serious concerns. Administrators in these districts had researched test taking, especially among younger children, and felt strongly that school-wide testing was not necessary and was not assessing the actual curriculum.

Dr. Herman Martinez, principal in San Luis, a largely Hispanic district, and Jane Martin, Staff Development Coordinator in Del Norte, which is partly Hispanic, used the columns as supporting material when they talked with community members about testing and when they approached their boards.

Within a month after testing, the two school districts revised their testing policies. In San Luis, tests will no longer be administered to K-3 students. In Del Norte, tests will only be administered in grades 4, 8, and 11.

Most parent reaction to testing, especially requesting that children be withheld from testing, was initiated by my columns. Parents (and even some administrators) learned that the state does not require testing in all grades. They were surprised to learn that they have a choice about having their children subjected to standardized testing. Parents also raised questions about how test scores affect their children's placement in special programs. Many were not satisfied with the ambiguous answers they usually received.

The doors for reform have opened. One parent has suggested forming a group, Parents Against Standardized Testing, to challenge districts on the use and abuse of tests. Teachers and parents now know we are not bound by the invisible chains of mandatory testing. We have a choice. ❑

NAEP Math Results Released

Results of the 1990 National Assessment of Educational Progress (NAEP) math test, including the first trial state comparisons, were released in June. *The State of Mathematics Achievement: Executive Summary* reported, "The mathematical skills of our nation's children are generally insufficient to cope with either on-the-job demands for problem solving or college expectations for mathematical literacy."

In the trial state comparisons for 8th grade math, the states with the most homogenous populations, no large cities and fewest people in poverty scored highest. As in the past, southern states averaged lower scores. Three of the four top-ranking states were among the states with the lowest portion of their students tracked by "ability grouping."

Students in "advantaged urban areas" scored highest on the NAEP math tests, those in disadvantaged urban schools scored lowest. By racial/ethnic group, Asian/Pacific Islanders scored highest, followed by whites, Hispanics, and then blacks, with sizable gaps between the first two groups and the second two. Only in grade 12 did a gender gap, favoring males, appear. Trend data comparing the 1990 assessment results with previous years will be reported in the fall.

For 12th grade students in academic programs, achievement was considered low -- slightly below 300 (on a 0-500 scale), with 300 reflecting a grasp of elementary geometry and algebra, which are commonly introduced by grade 7. The level of math understanding and problem-solving ability was viewed by NAEP as quite low, even for college-bound students.

Students and teachers both reported large amounts of testing in their classes, as well as curricula emphasizing heavy reliance on rote work with textbooks and worksheets. Students in low scoring states tended to be tested more frequently, and low-scoring students in general tended to be tested more often. Whether frequent testing contributes to low scores could not be determined from the data.

Other than one special sample, the exam remained predominantly multiple-choice. Most of the "open-ended" items that were included were not the sort of complex, multi-step problems with different solution paths that are being pioneered, for example, in the California Assessment Project. Rather, they were essentially multiple-choice items with the answer-options removed.

Reporting NAEP

The NAEP report analyzing math scores urged great caution in the interpretation of its data. For example, while large amounts of television watching correlated in general with lower test scores, the study could not be used to claim that increased TV watching causes lower scores. (In Washington, D.C. which scored below every state, there tended to be little difference in TV watching between high- and low-scoring students.)

Despite the cautions, the rush to reach conclusions based solely on test data without necessary contextual information still affected reporting by the media. One report, for example, claimed the NAEP results showed that states did not do better by following the new standards of the National Council of Teachers of Mathematics (NCTM). In fact, teachers around the nation had almost no opportunity to implement the new curriculum and instructional practices recommended by NCTM prior to the NAEP testing. Nor was the 1990 NAEP math test constructed to assess student abilities in ways that matched the standards. Reports such as these show that relying solely on test data to make educational decisions runs the risk of damaging effective programs.

Whether the NAEP results will be used to improve educational practices remains to be seen. California Superintendent William Honig used the release of the NAEP information to announce changes in California's curriculum that had already been planned. Though it has long been known that low-income, inner-city children tend to score low and that the South scores lower than the rest of the nation, this information has not been used to change education so as to provide equal opportunity to all children.

Math Standards Set

The National Assessment Governing Board (NAGB) voted in May to establish national standards for student achievement on the NAEP assessments in math in grades 4, 8 and 12. Although three panels of experts had declared the standard-setting process to be "flawed," NAGB adopted three levels ("basic," "proficient," "advanced") that will be applied to the 1990 NAEP math results. Rather than developing agreed-upon standards and then constructing a test to determine how well students measured up to the standards, NAEP standards were derived from the test after it had been administered.

A report on how well students did as compared with the standards will be released in late summer or early fall. ❑

Criteria for Assessment Systems Established

Dozens of education, civil rights and advocacy organizations have endorsed the National Forum on Assessment's "Criteria for Evaluation of Student Assessment Systems." Initiated by FairTest and co-chaired by Ruth Mitchell of the Council for Basic Education and FairTest's Monty Neill, the Criteria have been endorsed by organizations as diverse as the Council of Chief State School Officers, the Children's Defense Fund, the International Reading Association, the NAACP, and the National Education Association.

The eight criteria establish a solid foundation for what high-quality assessment systems should do. The Forum views the Criteria as applicable to all levels of assessment, from classroom to state or -- should it happen -- national.

The Forum criteria include:

- Defining standards for what students should know and be able to do before implementing tests;
- Designing assessment systems primarily to help educators and policy-makers improve instruction and student learning;
- Making sure assessment standards, tasks and uses are fair to all students and are understandable to users of assessment information;
- Reporting test results in the context of other relevant information, such as the resources devoted to education and types of educational programs;
- Involving teachers in designing and implementing an assessment system; and
- Ensuring that assessment systems are subject to continuous review and improvement.

Developed over a six-month period by representatives of several dozen organizations, the eight criteria are explained in detail in the Forum document.

▼ For a copy of the Criteria, a list of signers, and an endorsement form, write to the National Forum at FairTest. ❑

NAECS Recommends Eliminating Early Tests

The **National Association of Early Childhood Specialists in State Departments of Education** has recommended that standardized testing requirements for Chapter I be eliminated prior to grade 4. Currently, a norm-referenced standardized test must be administered to each Chapter I student each year beginning in second grade for schools to continue to receive federal funds for disadvantaged students. A position paper sent to federal Chapter I officials, state directors of Chapter I, chief state school officers, and heads of advocacy groups concludes that the standardized test requirements conflict with developmentally appropriate educational practices and perpetuate harmful practices such as ability grouping, also known as tracking.

▼ For copies of NAECS/SDE's "Chapter I Services and Early Childhood Education: Problem or Promise," contact Chalmer Moore, Jr., Illinois Board of Eduction, 100 North First St., S-100, Springfield, IL 62777. (*Education Week*, 6/5/91). ❑

Larry P. Narrowed

Larry P. v. Riles, the California case which barred the use of IQ tests for placing black children in special education classes, has been narrowed as a result of a challenge by the conservative Landmark Legal Foundation. In *Larry P.*, US District Court Judge Robert Peckham found IQ tests to be racially biased against black children. In 1986, Peckham issued a consent decree in which California agreed to a prohibition on all uses of IQ tests for assessing black children for placement in special education.

In *Crawford v. Honig*, several parents represented by Landmark argued that prohibiting IQ testing for black children amounted to discrimination, since whites and members of other minority groups could be tested. Landmark sought a complete reversal of the *Larry P.* consent decree.

Instead, Peckham allowed IQ tests to be administered to two black children for special education placement purposes. He did not allow any other children to be excepted from the consent decree. Rather, any parent who seeks to have a black child administered an IQ test to help decide placement in special education will also have to appeal through the court.

Public Advocates, the group which fought for the black children in *Larry P.*, urged Peckham to dismiss *Crawford*. They argued that while the tests discriminated against all minority children, blacks as a class have the right to protect themselves against the tests even if other minority groups have not done so.

"Judge Peckham's decision shows the weakness of remedies for only one race," said Armando Menocal of Public Advocates. "IQ tests should be banned for all children, not just blacks." ❑

K-12 Shorts

Indiana has dropped 1st and 11th grades from the state's standardized testing program. The state now tests in grades 2, 3, 6, 8 and 9. Low-scoring pupils can be required to attend summer remedial programs before entering the next grade. The Department of Education supported dropping the grade 1 test, but opposed ending the 11th grade test. Governor Bayh supported the measure, which passed both houses by 2-1 margins, but he has called for mandatory testing of high school seniors. (*Education Week* 5/22/91)

Students in some **Louisiana** schools have protested against classmates being denied the right to graduate because they did not obtain a passing score on the state-mandated Louisiana Education Assessment Program (LEAP). In Vidalia, 60 students signed a petition asking permission for an honor-roll classmate who failed the exit-exam by one point to be able to participate in graduation exercises.

Across the state; students who failed by close margins were denied a diploma, preventing some from attending the colleges to which they had been admitted. Students claimed that test-induced stress lowered their performance. Some reported signs of stress such as nosebleeds and ulcers, while one parent said her son was hospitalized for a "nervous condition." ❑

New York Approves New Assessments

A combination of portfolios of students' best work, teacher evaluations and standardized exams will replace most state-mandated multiple-choice testing programs in New York. The standardized exams will be only partly multiple-choice. The new assessments for grades 4, 8, and 12 are part of Commissioner of Education Thomas Sobol's "New Compact for Learning" approved by the Board of Regents in June. A Curriculum and Assessment Council is being created to provide guidelines for implementation of the assessments.

New York Campaign

The New York Campaign for Genuine Accountability in Education, the FairTest-initiated coalition working to ban standardized testing of young children, was invited to participate in a State Education Department (SED) meeting to discuss draft regulations on the testing of young children. The SED will present its recommendations to the Board of Regents, which is likely to consider them this fall. The Campaign did not agree with the SED draft, which essentially recommends that districts exercise caution in the use of standardized tests. The Campaign's position is that these tests are educationally harmful, so state regulations should prohibit or severely restrict their use. The Campaign will present its own proposal to the Regents. Campaign members are meeting with Regents in New York City, Binghamton, Syracuse, Rochester, Buffalo, and Westchester.

Ruth Messenger, Manhattan Borough President, has joined the list of endorsers of the Campaign's "Statement to Urge a Ban on Standardized Testing of Young Children." ❑

Civil Rights Bill

(Continued from page 16)

academic achievement are legally defensible."

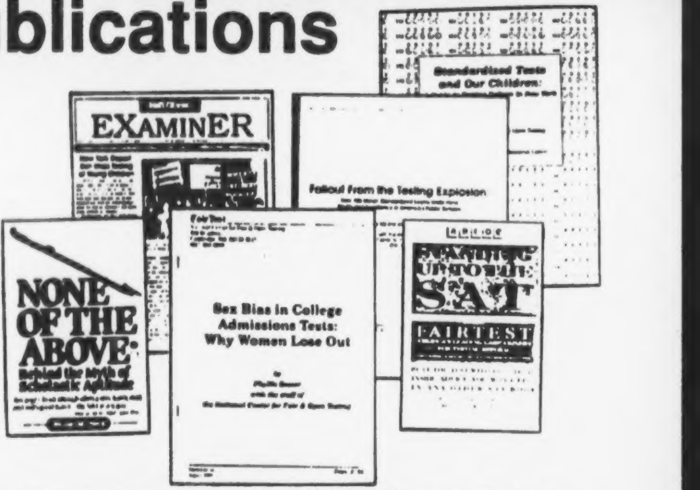
Kemp and his allies have also sought to use the Civil Rights Act debate to revise EEOC's Uniform Guidelines on Employment Selection Procedures, which bar tests that discriminate against women and minority workers unless the criteria are clearly job-related. Some analysts believe that President Bush's plan to make results from his proposed *America 2000* educational achievement tests available to prospective employees would violate current EEOC guidelines. The guidelines grew out of the 1971 *Griggs v. Duke Power* case that struck down personnel policies which act as "built-in headwinds" to minorities.

Sponsors of the Civil Rights Act have responded to the heated debate by adding provisions barring within-group scoring but requiring employment tests to be proven fair and valid for the job in question. The GATB is unlikely to meet this standard since its use for a wide variety of jobs is based on "validity generalization." That theory holds that one test can be used to predict performance on a variety of jobs without validity or fairness studies for each particular job.

As amended in the U.S. House, the civil rights bill also outlaws the use of hiring quotas, except when they are ordered by a court. Despite these compromises, the White House, many Republicans, and some Democrats remain opposed to the bill.

Fair-employment advocates continue to seek wording that can produce a veto-proof majority in both houses of Congress. But no matter how that debate is resolved, the once-obscure issues of employment testing are certain to remain at the center of national policy-makers' attention for months to come. ❑

FairTest Publications



EXAMINER
NONE OF THE ABOVE
Sex Bias in College Admissions Tests: Why Women Lose Out
Fallout From the Testing Explosion
Standardized Tests and Our Children
STANDING UP TO THE SAT
FAIRTEST

FairTest Examiner, published quarterly...$15.00/year (individuals), $25.00/year (institutions) $______

Back issues of *FairTest Examiner*, $4.00 each:
Vol____ No ____(All 14 for $40.00) $______

Standing Up to The SAT, by John Weiss, Barbara Beckwith and Bob Schaeffer (192 pp.)...$6.95 $______

Fallout From the Testing Explosion: How 100 Million Standardized Exams Undermine Equity and Excellence in America's Public Schools, by Noe Medina and Monty Neill (77 pp.)...$8.95 $______

Sex Bias in College Admissions Tests: Why Women Lose Out, by Phyllis Rosser and FairTest staff, 3rd ed. (60 pp.)...$7.95 $______

The Reign of ETS, by Allan Nairn and Ralph Nader (550 pp.)...$30.00 $______

Beyond Standardized Tests: Admissions Alternatives That Work, by Amy Allina with FairTest staff (18 pp.)...$5.50 $______

None of the Above, Behind the Myth of Scholastic Aptitude, by David Owen (327 pp.)...$7.95 $______

Standardized Tests and Our Children: A Guide to Testing Reform, by FairTest (32 pp.)... $3; 5 for $10; 10 for $15; 50 for $50; 100 for $75. English and Spanish editions can be combined for bulk prices.

copies in English #______
copies in Spanish #______
Total $______

Standardized Tests and Our Children: A Guide to Testing Reform in New York, by FairTest & NYPIRC (32 pp.)...same prices as above. *Single* copies from FairTest. Bulk copies available from: NYPIRC, 184 Washington Ave., Albany, NY 12210

English #______
Spanish #______
Total $______

Contributions to FairTest

I agree there's no better way to be part of the testing reform movement than supporting FairTest! Enclosed is my tax-deductible contribution!*

☐ **FairTest Associate** ($25.00--includes a free subscription to the *FairTest Examiner*)

☐ **FairTest Sponsor** ($50.00--includes *Examiner* subscription and a FairTest publication of your choice, excluding *Reign of ETS*)

☐ **FairTest Sustainer** ($100.00 or more--includes above, plus a 20% discount on publications)

☐ **Other** ______

Publications Total $______
Contributions Total $______
Check Total $______

Name ________________
Address ________________
City ________ State ____ Zip ______

Please make checks payable to FairTest and mail to: FairTest, 342 Broadway, Cambridge, MA 02139

*Contributions to FairTest are fully tax deductible, except for the value of publications received.

Testing "Red Herring" Blocks Civil Rights Bill

Passage of the federal "Civil Rights and Women's Equity in Employment Act" has bogged down in a partisan dispute over the use of test scores to make hiring and promotional decisions. The bill is designed to overturn several recent U.S. Supreme Court decisions which made it more difficult for women and minorities to win job discrimination cases.

A major item of contention in the emotional debate has been a procedure called "within-group scoring" which is applied to the General Aptitude Test Battery (GATB) designed by the U.S. Department of Labor. The GATB is used by thirty-eight state employment offices to make job referrals. Labelled "race norming" by its opponents, the ten-year-old practice allows black and Hispanic applicants to be referred to jobs on the basis of their test score ranking within their own racial groups, not the entire population of job-seekers. Because blacks and Hispanics have average lower scores than whites, they may be referred (and hired) before whites with higher scores. This, opponents say, creates "quotas" which unfairly exclude "better qualified" white workers.

"Within-group scoring" was specifically endorsed by a 1989 National Research Council (NRC) report *Fairness in Employment Testing* which concluded that ". . . the disproportionate impact of selection error provides scientific grounds for the adjustment of minority scores so that able minority workers have approximately the same chances of referral as able majority workers."

The NRC study also criticized the GATB's inaccuracy saying, "the predictive power of the test battery is not so strong that the GATB should become the sole means of filling all job orders." Following the NRC report, former Secretary of Labor Elizabeth Dole proposed suspending the GATB pending a thorough study of all issues relating to its use (see *Examiner*, Fall, 1990). No final action has yet been taken on her recommendation.

> ***"The predictive power of the test battery is not so strong that the GATB should become the sole means of filling all job orders."***
>
> ***-National Research Council***

Other test critics, led by the Lawyers Committee for Civil Rights Under Law, have demonstrated that the GATB is neither job-related nor fair. They also note that several tests in the battery are highly speeded and reward sophisticated guessing strategies. Thus, the GATB identifies good test-takers, not necessarily those who would make the best employees. Opponents of within-group scoring confuse high test scores with real job qualifications.

Opposition to the civil rights bill and within-group scoring has been led by Evan J. Kemp, Jr., Chairman of the Equal Employment Opportunity Commission (EEOC), and other White House advisors. Kemp has written that passage of this civil rights bill would make it "extremely difficult for employers to show that use of educational credentials and objective measures of

(Continued on page 14)

FairTest®
National Center for Fair & Open Testing
342 Broadway
Cambridge, MA 02139-1802
(617) 864-4810
(617) 497-2224 (FAX)

Non-Profit
U.S. Postage
PAID
Cambridge, MA
Permit No. 53278

ADDRESS CORRECTION REQUESTED

***Do you have an asterisk on your label?** If you do, this is your last issue, so **renew now!**

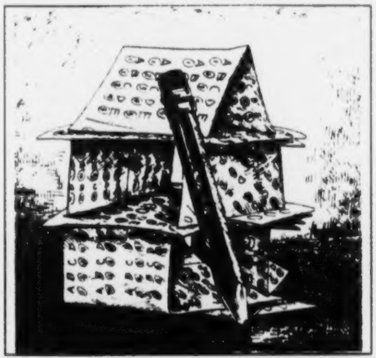
Moving? Return your label with new address

FairTest
The National Center for Fair & Open Testing
Examiner
Volume 5, No. 4 Fall 1991 ISSN #0898-2511

NEW FAIRTEST REPORT

SAT Coaching: Reality vs. Cover-up

A new FairTest report concludes that the Scholastic Aptitude Test (SAT) is extremely coachable, contrary to the claims of its manufacturers. *The SAT Coaching Coverup* reviews research studies showing that students taking commercial test preparation courses can expect to increase their SAT scores an average of 100 points or more.


Kurt Mitchell, reprinted from *Catalyst*

The fact that coaching works undermines one of the basic premises of the SAT, that it measures skills learned over a long period of time. The SAT's "coachability" benefits those who can afford expensive test preparation courses, often costing $500 or more, while harming those who are already unfairly disadvantaged by the test's biases — minorities, low-income students and those whose first language is not English. This income-related bias gives affluent students an unfair edge in winning scholarships and admission to colleges which rely heavily upon test scores. The College Board's repeated denials of coaching's overall effectiveness are both false and potentially damaging to students who believe them.

Many studies reviewed in the report showed large score gains from coaching. But just as it took only one flight at Kitty Hawk to prove wrong those who said "humans can't fly," only one effective test preparation program is needed to disprove the test-makers' nay-saying. The "Wright brothers flight" of coaching studies was conducted by researcher John Zuman, who found average increases of 110 points as the result of test preparation in his 1988 study, "The Effectiveness of Special Preparation for the SAT: An Evaluation of a Commercial Coaching School." Zuman, then an Ed.D. candidate at Harvard University, was advised by George Madaus, a distinguished psychometrician and former president of the National Council of Educational Measurement and Evaluation. Zuman concluded:

CONTINUED ON PAGE TWO

Congress Begins to Focus on Testing

As Bush administration efforts to establish national testing accelerate, members of Congress are far from unified on this issue. While Sen. Kennedy (D-MA), chair of the Labor and Education Committee, has indicated he has reservations about national testing, Sen. Pell (D-RI), chair of the Subcommittee on Education, is pushing for a national exam.

Last spring, Pell introduced legislation for a national test or series of tests. Pell now seeks to include his proposal as part of this fall's reauthorization of the Department of Education's Office of Educational Research and Improvement (OERI).

The Administration wants to use the OERI reauthorization as a vehicle for removing the prohibitions on the use of National Assessment of Educational Progress (NAEP) data to rank schools and on the use of NAEP test items in constructing state, district or school-level tests. This proposal appears to lack Senate or House

CONTINUED ON PAGE EIGHT

TABLE OF CONTENTS

3 University Testing
SAT Coaching: Reality vs. Cover-up 1
New SAT and ACT Scores 4
College Board Stops Listing Cut Scores 5
ACT Fact Sheet 6

8 K-12 Testing
Congress Begins to Focus on Testing 1
Campaign Sends Letter to Council 9
Controversy Erupts Over NAEP 10
Connecticut's Performance Assessments 12
Controversial Oregon Reforms Pose Questions for Future of Assessment 13

16 Employment Testing
More Hurdles for Prospective Teachers 16

TM017184

FairTest
National Center for
Fair & Open Testing
342 Broadway, Cambridge, MA 02139-1802
(617) 864-4810 •(617) 497-2224 (fax)

STAFF:

EXECUTIVE DIRECTOR
Cinthia Schuman, Esq.
ASSOCIATE DIRECTOR
Monty Neill, Ed.D.
UNIVERSITY TESTS DIRECTOR
Sarah Stockwell
FIELD ORGANIZER
Veda Wright
PUBLIC EDUCATION DIRECTOR
Bob Schaeffer
PUBLICATIONS/OFFICE MANAGER
Jamie Souder

BOARD OF DIRECTORS:

Ron Ellis, Esq., NAACP Legal Defense and Educational Fund; **Dr. Pamela George**, North Carolina Central University; **Dr. Deborah Meier**, Central Park East Secondary School; **Dr. Susana Navarro,** Southwest Center for Academic Excellence, University of Texas at El Paso; **Dr. Vito Perrone,** Harvard Graduate School of Education; **William Robinson, Esq.**, Dean, D.C. School of Law; **Bob Schaeffer**, Treasurer; **Steve Solomon, Esq.**, Proskauer, Rose, Goetz & Mendelsohn; **Chuck Stone**, University of North Carolina, Chapel Hill; **Ron Vera, Esq.**, Best, Best & Krieger, Board Chair; **John Weiss**, Public Policy Initiatives, Board President.

NATIONAL POLICY PANEL:

Dr. Leonard Beckum, Duke University; **Dr. Beverly Cole**, NAACP; **Dr. Harold Dent**, Association of Black Psychologists; **Arnold Fege**, National PTA; **Dr. Howard Gardner**, Harvard University; **Dr. Asa Hilliard**, Georgia State University; **Dr. Lucinda Lee Katz**, University of Chicago Laboratory Schools; **U.S. Rep. Matthew Martinez**, (D-CA); **Aurelio Montemayor**, Intercultural Development Research Association; **Harry Pachon**, National Association of Latino Elected and Appointed Officials; **Alfredo de los Santos**, Maricopa Community Colleges; **Eleanor Smeal**, Fund for a Feminist Majority; **Theodore Shaw, Esq.**, University of Michigan Law School; **Dr. Gail Thomas**, Texas A&M University; **U.S. Rep. Ted Weiss**, (D-NY).

Funding for FairTest's work comes from the **Golden Rule Insurance Co., Ford Foundation, Rockefeller Family Fund, Philip M. Stern Family Fund, Aaron Diamond Foundation, Edward Hazen Foundation, Lilly Endowment, New York Community Trust, and many individual contributors.**

PAGE 1 STORY CONTINUED

New FairTest Report

If in fact coaching can substantially increase students' SAT scores, then it is the wealthier families, who already have an advantage in the college admissions process, who benefit further by being able to afford to take the expensive but effective coaching programs and increase their SAT scores... If coached students have inflated SAT scores, then colleges may well be turning away some more talented students only because those students did not attend an effective SAT preparation course, and may instead be admitting students from wealthier homes.

The FairTest report cites many other studies which document the effectiveness of coaching programs, including school-based workshops, computer software, and special programs targeted toward minority students. Score increases broadly agree with Zuman's findings, ranging from 50 to 135 points.

The SAT Coaching Coverup shows that several studies which claim coaching is ineffective are marred by researcher conflicts-of-interest and flawed methodology. For example, the best-publicized of such studies, "Coaching: Does it Pay? Not for Harvard Students," (*College Board Review*, Summer 1988) was conducted by Dean Whitla, a College Board Trustee and paid consultant to the Educational Testing Service, who is director of the Office of Instructional Research and Evaluation at Harvard University. In 1962, Whitla wrote that if coaching worked, "'the validity and essential purposes for which the test was constructed' would be challenged" (as quoted in David Owen's *None of the Above*). Not surprisingly, in his study Whitla finds that "coaching did not raise the SAT scores of Harvard students in any significant way."

In addition to his conflict of interest, Whitla's method of data collection raises questions as to the accuracy of his conclusions. First-year students at Harvard were asked, at the time of registration, to tell college officials the extent to which they had engaged in test prepara-

CONTINUED ON NEXT PAGE

Fresh Challenge for FairTest Staffer

Laura Pierce, FairTest's adminstrative assistant, left in October to become the development coordinator of the *Gay Community News*, a national newsweekly. Laura's skills and presence are missed by everyone. We wish her the best of luck in her new career endeavors. ❑

A New Look for the FairTest *Examiner*

It's been four years since we published the first issue of the FairTest *Examiner*. This is an appropriate anniversary, we thought, for updating the look of this quarterly source of information about standardized testing reform.

Key to the Examiner's re-design are clearer separation of the three significant areas in which FairTest works— K-12, University and Employment Testing.

We hope you like our new look! ❑

SAT Coaching: Reality vs. Cover-up

tion activities. Many students may have felt they would be incriminating themselves by revealing how much they were coached. Whitla relies on these questionable data to conclude that his results provide confirmation of coaching's ineffectiveness "far beyond our particular study and sample" — a statement which is not substantiated.

Other studies which found low effects for coaching courses have similar problems. The FairTest report shows how they relied on direct adaptations of Whitla's survey technique to overgeneralize from high-scoring populations of private school students.

While test-makers have had to respond to the growing evidence, they still deny coaching's overall effectiveness. Where once the College Board denied that coaching could increase students' scores, *Taking the SAT*, their free booklet, now tells students "We are not saying [coaching courses] don't work. We just can't say with certainty whether a particular program or activity will work for you." Such cautions ignore the fact that dramatic score increases occur *on average*. Thus, the College Board still deliberately underplays the effectiveness of coaching to students.

FairTest's report concludes that the "new" SAT, scheduled to be implemented in the Spring of 1994, is likely to be as coachable as the current version. Initial studies indicate that one of the most coachable item types, reading comprehension, will have even greater weight in the new test. Coaching companies typically advise test-takers to read the answer choices on reading comprehension items first (the opposite of the directions given on the test) and then locate any information they may need to answer the question. Research by Dr. Stuart Katz at the University of Georgia has confirmed that this technique works. In fact, Katz found that students can choose the correct answer to about 43% of the questions on new SAT reading comprehension items without ever reading the passage. Furthermore, Katz' data shows that new SAT reading comprehension items are even easier than the current ones, which students answered correctly at a rate of 36%.

From the Inside...Out

Lewis Pike, a former ETS researcher who was fired shortly after the publication of a study that showed coaching worked, had this to say about test coaching:

"One of my main concerns ...was what happens to real people out there. When people who have money debate whether to pay for coaching, they say, "Well, it can't hurt and it might help, so we might as well shell out the bucks." But if you're poor, and don't have the money, and someone quotes ETS or the College Board saying that coaching doesn't work, then you kind of breathe a sigh of relief and say, "Well, it doesn't work anyway, and thank God, because where would we get the three hundred bucks?" These people get what I call double jeopardy; that is, they're getting shafted in the schooling they receive and getting shafted by the SAT. You're never going to eradicate all of that, but if you pretend that coaching doesn't help, you're being irresponsible and you're adding to the problem instead of helping it."

— from *None of the Above*, by David Owen (Houghton Mifflin, 1985.)

Another current item type, antonyms — said by SAT coaches to be the most difficult verbal item type to coach — will be dropped from the new SAT. Coaching companies say they expect to have no trouble preparing students for any part of the new SAT, including the new "grid-in" math items.

Based on this new report, ETS and the College Board have an obligation to tell test-takers the truth — that coaching works — and to inform colleges that test preparation may seriously affect the scores students report. If the test-makers still believe that coaching is ineffective, despite the overwhelming evidence, FairTest challenges them to one "final" study, conducted by an independent blue-ribbon panel, to resolve any lingering doubt.

✎ Copies of *The SAT Coaching Cover-up* are available for $ 7.95 postpaid from FairTest, 342 Broadway, Cambridge MA 02139. ❑

Annual SAT and ACT Scores: Unchanged and Overinterpreted

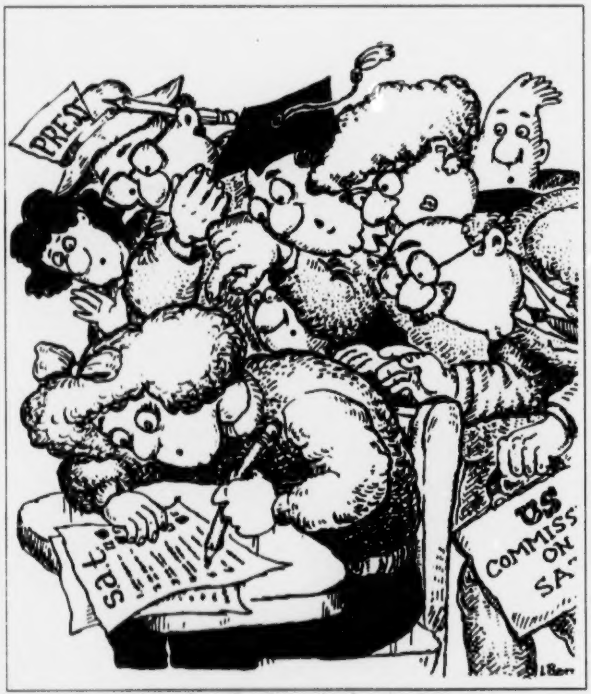
The College Board's annual release of SAT averages was accompanied by the usual overinterpretation of minor score variations: this year's continuing small decline in the SAT-Verbal score brought warnings of a "verbal free fall" from CEEB President Donald Stewart.

In fact, average scores on the SAT dropped only two points from last year to 422 and 474 on the Verbal and the Math section respectively. Long-term trends show that Verbal scores have fluctuated within a 10 point range for the past 15 years. Yet the College Board claimed this year's scores indicated the need for "drastic measures" to meet President Bush's educational goals.

The average ACT composite score was 20.6 —unchanged from last year. Over five years, ACT scores have declined only .2 points (when old scores are converted to the new "Enhanced" ACT scale).

Scores of minority groups have remained relatively stable, still far behind white averages, and males still score half a point higher than women. As in the past, ACT's analysis emphasized comparisons between students who had taken a "core" of college preparatory coursework and those who had not, while downplaying race, income and gender differences.

In public discussion of score changes, important questions all too often remained unaddressed, such as what SAT or ACT scores actually reveal about U.S. educational achievement or why the College Board releases score averages by state despite its own recognition that comparisons of state SAT scores are "invalid and strongly discouraged...."



Loel Barr from *The Chronicle of Higher Education*

Critics have long dismissed changes in annual score averages as meaningless, in part because the population taking the tests varies. Some researchers even claim that annual test score averages reveal a fairly steady level of test-taking achievement over many decades.

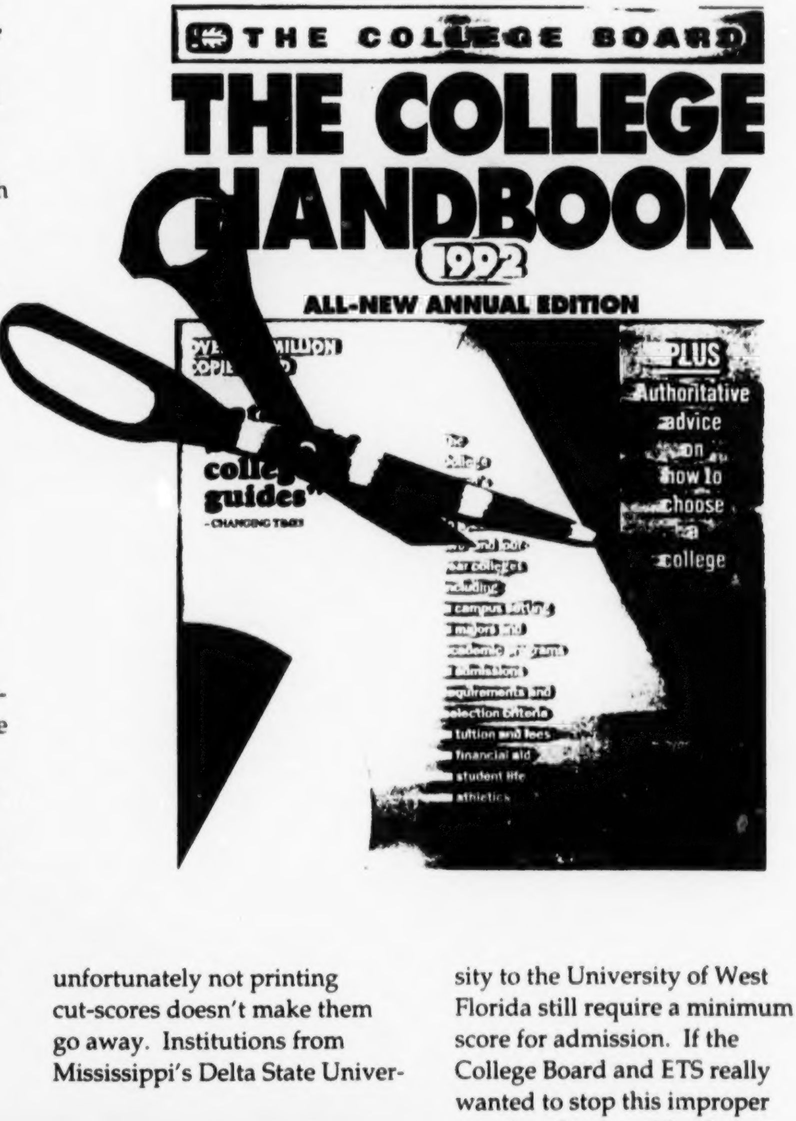
Comparisons of state averages are also misleading both because of the varying proportion of students taking the test and differences within the student populations. For instance, an increase in the number of students for whom English is a second language can alter a state's average score without revealing anything about the quality of its education. The result is that states like South Carolina, with 58% of its diverse student population taking the SAT, is compared unfavorably to Iowa, which has a very small proportion of non-white students and where only five percent of students, generally those bound for highly competitive schools, take the SAT.

The College Board would do a service to educational reform by refusing, as does ACT, to release annual score averages by state. Such data reveal almost nothing about student achievement, but have harmful effects on school, district and state efforts to enact real educational reforms.

A FairTest statement responding to the test score announcements concluded, "The SAT and ACT are models for the kinds of destructive tests that are all too common in U.S. education: multiple-choice, norm-referenced instruments which are divorced from real curriculum, and fail to measure students' ability to think, solve problems, and use knowledge." ❑

College Board Drops *Handbook* Cut Scores

Responding to criticism from FairTest, the College Board has stopped reporting "cut-off scores" in its 1992 version of *The College Handbook.* The annual *Handbook* profiles 3200 two- and four-year colleges in the U.S., and is widely relied upon by college-bound students and high school guidance counselors for information.



The use of SAT cut-off scores — or minimum test score requirements — for college admission violates the College Board's *Guidelines on the Uses of College Board Test Scores and Related Data,* which warns admissions officials to "guard against using minimum test scores unless used in conjunction with secondary school performance and unless properly validated." One year ago, FairTest pointed out the discrepancy between this policy and the data in the *College Handbook.* When asked to stop publishing score cut-offs, the College Board's initial reaction was that it could not do anything about this problem since "these are institutional decisions, they are not College Board decisions...."

While it is a step forward that the College Board addressed this contradiction, unfortunately not printing cut-scores doesn't make them go away. Institutions from Mississippi's Delta State University to the University of West Florida still require a minimum score for admission. If the College Board and ETS really wanted to stop this improper practice, they would refuse to send SAT scores to any school which uses cut scores.

FairTest has long contended that the use of cut-off scores magnifies the unfair cultural bias and imprecision of the SAT. The test's 68 point margin of error means that any use of a minimum score has a particularly harmful effect on minorities and women, whose scores generally fall closer to the cut-off. ☐

NEWS FLASHES

Wheaton College, in Norton, Massachusetts, has announced that it will make test scores optional for college admission. The admissions office intends to focus on the applicant's high school academic record and commitment to life-long goals.

◆ ◆ ◆

The use of a cut-off score on the ACT for entrance into Alabama's teacher education program has been struck down by a federal court. For fifteen years, the Alabama State Board of Education required a minimum score of 17 on the ACT. The successful plaintiffs showed, among other things, the devastating impact the rigid cut-off score had on African American teacher candidates.

The "Enhanced" ACT: Still Biased,

More than a million high school students take the American College Testing Assessment (ACT) each year. Like the Scholastic Aptitude Test (SAT), the ACT is a standardized multiple-choice test meant to predict first-year college grades. While the SAT predominates in the Northeast and the West Coast, the ACT is more common in the Midwest and the South.

A new test?

In the fall of 1989, the American College Testing Program revised its test to be "responsive to changes that have occurred in high school curricula..." The new version consists of four individual tests: English, Math, Reading, and Science Reasoning. The score report for the "Enhanced ACT Assessment," as the new test is known, includes a composite score ranging from 1-36, a score for each individual test, and, for the first time, subscores. Though ACT modified the subject areas and added subscores, "all other components of the ACT Assessment remain essentially unchanged."

Unfortunately ACT did not address the long-standing problems of bias, inaccuracy, coachability and misuse in its test.

The ACT is biased

Race, class and gender biases give white, affluent, and male test-takers an unfair edge.

Score gaps between groups on the ACT cannot be explained away solely by differences in educational opportunity. According to ACT research, when all factors are equal, such as coursework, grades and family income, whites outscore all other groups — including Asian Americans. If the ACT were not biased, Asian Americans, who take more academic courses than any other group, and rank first when scores are not adjusted for such factors, would likely score even higher. Moreover, boys score higher than girls across all races, despite boys' lower grades in high school and college. ACT has performed few studies of score differences in its test, making it difficult to pinpoint the sources of bias. Here are a few likely candidates:

Biased format: Research shows that a multiple-choice format favors males over females. Guessing, a risk males are more likely to take, is rewarded. Since multiple-choice items do not allow for shades of meaning they work against females' more complex thinking style.

ACT Score Averages

Family Income	Score
$60,000 and over	22.4
$50,000 to $59,999	21.7
$42,000 to $49,999	21.4
$36,000 to $41,999	21.0
$30,000 to $35,999	20.7
$24,000 to $29,999	20.4
$18,000 to $23,999	19.9
$12,000 to $17,999	19.4
$ 6,000 to $11,999	18.6
less than $6,000	17.3

Race/Ethnicity	Score
Asian American	21.6
White	21.3
Puerto Rican/Hispanic	19.3
Mexian American	18.4
Native American	18.2
African American	17.0

Gender	Score
Male	20.9
Female	20.4

Average scores from *The ACT High School Profile Report*, for high school graduating class of 1991.

Biased language: Idiomatic terms such as "thumb its big nose at," and "straight from the horse's mouth," may not be familiar to many test-takers, particularly those whose first language isn't English, causing them to choose wrong answers.

Biased question context: Studies show test-takers do better on questions they find interesting, or that are set in familiar situations. Many more English and Reading ACT passages cover topics that are likely to be more familiar to whites and males than to minorities and to females.

One reason for this is that people of color and women are seldom featured on the ACT. Of four publically-disclosed tests, one did not include a single reference to a person of color and two mentioned only one. Similarly, men appeared in items five times more often than women. When people of color and females are mentioned, they rarely have the status of whites and males. For example, one test featured 21 white males, including famous scientists, politicians and artists. The single minority group member was a nameless Japanese tea-master. The five females included a "she," and four characters in a fiction passage.

The ACT is inaccurate

ACT scores do not predict freshman grades effectively.

Inaccurate, Coachable, and Misused

Even the test-maker admits that high school grades predict college grades better than ACT scores do. In fact, adding the ACT to the high school record does not significantly improve predictions.

The ACT regularly underestimates the abilities of females, who earn higher grades than males in high school and in college. Females receive higher grades than ACT scores predict in all courses of study, including pre-med and business. Recognizing the problem, the Massachusetts Institute of Technology routinely admits females with lower math scores because they find women perform as well as men.

ACT scores are imprecise

According to ACT, the individual tests have large margins of error on the 1-36 point scale: 1.50 points in English and Math, 2.21 in Reading, and 2.04 in Scientific Reasoning.

The ACT is coachable

ACT itself sells coaching products. Yet ACT insists that "for students who have not studied the content or grasped it... short term review is not likely to be of much benefit." While ACT acknowledges that familiarity and test-taking skills can affect a score, it also maintains that the descriptions and sample items included in the registration packet address them. If that claim is true, why is ACT selling "Official Guides" for $12.95 and videotapes containing "effective strategies specific to each test on the Enhanced ACT Assessment" for $79.00? The National Association of Secondary School Principals also believes that the ACT is coachable; they promote an extensive line of ACT test preparation materials, ranging from workbooks to computer disks.

The ACT is misused

Cut-off scores on the ACT unfairly deny access to education and scholarships.

The ACT's flaws have serious consequences. Despite its inaccuracies, biases, and coachability, ACT cut-off scores are regularly used to determine entrance into schools and allocate scholarships. A single point can decide whether a student is admitted or receives needed funds. Though these misuses violate ACT guidelines, ACT has done nothing to stop them. ACT has the responsibility and the power to protect students from testing abuse by refusing to send scores to colleges, scholarship agencies, and educational systems which misuse their product.

ACT and Athletics

The National Collegiate Athletic Association's "Prop 48" denies athletic eligibility to students who do not score at least 18 on the ACT (or 700 on the SAT), no matter how good their high school grades are. Prop. 48 is not only ineffective at screening out unable students, but it discriminates against blacks, who are disproportionately denied eligibility. A study at the University of Michigan reviewed the pre-Prop. 48 records of students who would have been barred from competition due to low test scores. Eighty-six percent of these student-athletes successfully completed their first year and then graduated at rates higher than the general student population.

Education denied: In Mississippi, ACT test scores are used as a tool of segregation in state colleges and universities. In 1962 Mississippi's higher education board instituted a minimum ACT score for admission that was eight points above the Black score average. According to a federal court, this was done "soon after the court ordered admission of James Meredith to the University of Mississippi because it deterred black enrollment." The same court noted that today, because of Mississippi's adherence to a minimum score on the ACT, "The integration of the historically white universities is not invited," in Mississippi's state system.

Even if they earn high college grades, students cannot enter a public graduate school of education in Alabama without scoring at least 17 on the Enhanced ACT.

Scholarships denied: In Louisiana, high achieving students can have their tuition and fees waived for up to five years at state colleges and universities if they score above an ACT cut-off of 20. This minimum score is close to the average for white students in Louisiana, but more than three points above the African American average. A number of other states are starting similar programs.

✎ For a copy of this fact sheet in handout form, send a SASE to "ACT Fact Sheet" to FairTest. ❑

PAGE 1 STORY CONTINUED

Congress Begins to Focus on Testing

support. A prohibition on reporting individual student scores would remain. The Administration also seeks to expand the NAEP trial state-by-state comparisons that began last year. This proposal is widely supported.

The reauthorization legislation is likely to include an expanded OERI. One idea is to create a series of institutes to direct research, one of which would cover curriculum and assessment. Concern has already arisen that such an institute could proceed to develop a national test, perhaps through a contract to a private corporation, regardless of any action by Congress. In the past, critics have charged that the Administration has spent research money in ways not authorized by Congress. This year, when Congress appropriated funds for OERI, use of the money was carefully fenced-in to prevent it from being used as a Departmental slush fund (see *Examiner*, Spring 1991).

Other Legislation

Taking a different approach is H.R.3320, introduced by Reps. Kildee (D-MI) and Goodling (R-PA). Their proposal would include funding to develop new assessments as part of state and local education reform efforts. One quarter of the money could be used at the state level including for development of new assessments tied to new standards and curriculum; three quarters would be used for reform activities at the local level, including staff development. Organizations that support assessment reform but oppose a national test at this time agree with the thrust of the bill, but think it still needs improvement. However, the legislation was moved through the full Education and Labor Committee very quickly, providing little time for public comment or amendment.

Also in the Senate is S.2, an elementary and secondary education bill introduced by Sen. Kennedy. S.2 codifies the six education goals developed by the Administration and the Governors. It also establishes a National Council on Education Goals to monitor progress and issue an annual "Report Card."

Experts Urge Caution on National Testing

The American Psychological Association (APA), American Educational Research Association, and National Council on Measurement in Education have called on "Congress, President Bush and education policymakers to slow down the rush to implement national achievement tests and a national assessment system," warning that "the rush to establish [them] could ultimately damage the nation's education system."

The three groups are the joint authors of the *Standards for Educational and Psychological Testing*. The organizations have raised "serious technical issues" with the testing proposals in the President's America 2000 plan that should be addressed "prior to implementing a national assessment system." Their concerns substantially overlap those of the FairTest-initiated Campaign for Genuine Accountability.

✎ For a copy of their statement, contact APA, 1200 17th St. NW, Washington, DC 20036. ❑

Finally, S.1790, introduced by Sens. Kennedy (D-Mass) and Hatfield (R-OR) and by Reps. Gephardt (D-MO) and Regula (R-OH), is labeled the "High Skills, Competitive Workforce Act of 1991." Modeled on proposals found in the *America's Choice: High Skills or Low Wages* report from the National Center on Education and the Economy (NCEE), this bill addresses itself to "fostering a high skills, high quality, high performance workforce," in part through developing "standards and certifications of mastery" in various occupations, and "pilot projects for assessments in specific subject areas" on a statewide or multi-state basis. It calls for performance assessments, student projects and portfolios.

In filing the bill, Hatfield specifically noted the Oregon education reform act that relies heavily on the NCEE report (see article, page 13). S.1790 also appears to rely on the work of the Department of Labor Secretary's Commission on Acquiring Necessary Skills (SCANS). It authorizes creation of a national board to oversee development of voluntary national educational standards for various occupations that could lead to students' obtaining certificates in those fields. The board would oversee development of curricula and assessments to match the national occupational standards. ❑

Campaign Calls on Council to Reverse National Testing Decision

Concerned that "the National Council on Educational Standards and Testing has put the cart of testing before the horse of education reform," the FairTest-initiated Campaign for Genuine Accountability in Education has called on the Council to reconsider whether the U.S. should adopt a national testing system.

Congress established the Council in collaboration with the National Educational Goals Panel and charged it with determining the desirability and feasibility of national standards and tests. In August, the Council voted "tentatively" to support both.

In an open letter delivered to members of the Council for their Oct. 21 meeting, the Campaign wrote, "The Council's . . . fundamental task of determining desirability of standards and tests has not been satisfied." The letter explained that the Council had not sought input from the many civil rights and education groups that have reservations about national testing. "Haste is the enemy of quality in the development of performance assessment systems," the letter added. It urged the Council to hold a series of hearings across the nation to address the many significant, but unexamined, issues.

Even members of the Council have expressed concern that it has acted too rapidly and gone beyond the scope of its mandate. Last month, its four Congressional members — Sens. Bingaman (D-NM) and Hatch (R-UT) and Reps. Kildee (D-MI) and Goodling (R-PA) — signed a letter protesting the Council's haste.

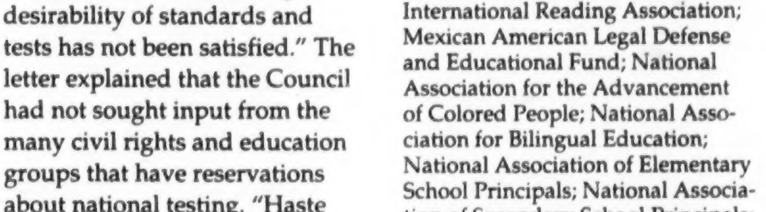
The Campaign's letter reasserted its positions that national testing without other fundamental education reforms will hurt, not help, US education; and that equity concerns must be addressed so that all children will be provided a fair chance to pass any exams.

* * *

Other signers of the letter included: American Association of School Administrators; ASPIRA, Inc.; Center for Women Policy Studies; International Reading Association; Mexican American Legal Defense and Educational Fund; National Association for the Advancement of Colored People; National Association for Bilingual Education; National Association of Elementary School Principals; National Association of Secondary School Principals; National Coalition of Title I/Chapter I Parents; National Council for the Social Studies; National Council of La Raza; National PTA; Puerto Rican Legal Defense and Education Fund; and Southern Association on Children Under Six.

✎ For a complete copy of the open letter, send an SASE to "Council Letter" at FairTest. ❑

By permission of Mel Lazarus and Creators Syndicate


TODAY'S TEST WILL BE MULTIPLE CHOICE.
I BELIEVE THAT MEANS WE CAN TAKE IT OR NOT...
3-6
MEL LAZARUS

NEA Favors Testing Moratorium

At its 1991 Representative Assembly, the National Education Association (NEA) voted to endorse a "Moratorium on Standardized Testing." The resolution reads:

"The Representative Assembly directs the NEA to utilize existing publications and form coalitions to call for and secure a moratorium on standardized testing of students in pre-kindergarten through grade 12 unless such testing is being used for diagnostic purposes or special needs screening. The moratorium shall continue until such time as the validity and detrimental effects of standardized testing on children have been determined." ❑

K-12 SHORTS

The National Center for Research on Evaluation, Standards, and Student Testing (CRESST) is a clearinghouse on research and development of assessment. CRESST serves educational policymakers, practitioners, and the public. The Center offers an extensive research database of over 340 assessment reports, monographs, and papers. For information or technical assistance, write or call CRESST/UCLA, 145 Moore Hall, 405 Hilgard Avenue, Los Angeles, CA 90024-1522, (310) 206-1532.

* * *

Growing Without Schooling has compiled a directory of articles from their past 30 issues that could be useful to people working to challenge testing requirements, particularly homeschoolers. For a list of the articles and the issues in which they appear, send a SASE to GWS, 2269 Massachusetts Ave., Cambridge, MA 02140. ❑

Controversy Erupts over NAEP Achievement Levels

The achievement levels on the 1990 National Assessment of Educational Progress (NAEP) mathematics exams, released with great fanfare on September 30 by the National Assessment Governing Board (NAGB), are based on a process that its own academic evaluators found severely flawed. ("Basic," "proficient" and "advanced" levels were set on the exams to define what students should know in math in grades 4, 8 and 12.)

The implicit charge in the evaluation, that the achievement levels-setting process was driven by a political agenda, raises questions about the credibility of NAGB and the claims made on the basis of NAEP test results.

NAGB "aborted its plan for an independent evaluation of its achievement-levels project by three senior evaluators when it looked as though the results were strongly critical," charged a letter by Richard Jaeger, one of three members of a committee appointed by NAGB to review the process used to set the achievement levels.

The rejection of the report, wrote Jaeger, was accompanied by a "gag order" and included an "unconscionable attack" on the authors. Copies of the report, Jaeger's letter, and comments on the report from NAGB are circulating unofficially. NAGB's action has caused a controversy in Washington. Congressional leaders have asked the General Accounting Office to investigate.

Margot Booth, from *Education Week*

The Report

The evaluation committee of Daniel Stufflebeam, Michael Scriven, and Jaeger concluded that the achievement levels did not warrant "*unqualified* release" because the methodology used to set them was not sound. Use of the results as a "benchmark" for future exams "would likely misinform the public."

If this evaluation is correct, the achievement results widely reported in late September — fewer than two-thirds of US students attained the "basic" level appropriate to their grade, fewer than one fifth are "proficient," and less than 4% could do grade-level advanced work— are suspect. However, the evaluators said, the information does support the conclusion that math achievement is "grossly deficient." The committee urged NAGB to redesign the levels-setting procedures, rather than release the achievement report.

Politics and Evaluation

The evaluation committee also asked Congress to modify the composition of NAGB to include more measurement experts. A change in NAGB composition is important, said the committee, because the rush to set achievement levels was politically influenced.

For example, the NAEP math levels were set based on tests that had already been administered. Even before the levels-setting process commenced, many experts told NAGB that decisions about what students should know and be able to do should be made before exams are constructed, not on the basis of existing tests. Moreover, the math exams were criticized as an inadequate sample of mathematical knowledge. NAGB pushed ahead, however, ignoring the cautions and criticisms.

Although the draft of the evaluator's report was favorably reviewed by many other prominent psychometricians, it was rejected by NAGB. According to Jaeger, NAGB also blocked dissemination of a companion evaluation on the adequacy of the levels themselves, *The Validity and Credibility of the Achievement Levels*, authored by another team of respected psychometricians, Robert Linn, Daniel Koretz, Eva Baker and Leigh Burstein. This report is even more critical of the achievement levels than the Stufflebeam report, stating that the levels were not credible and should not be used publicly. After limited circulation of a draft, said Jaeger, the Linn panel's

CONTINUED ON NEXT PAGE

Controversy Erupts Over NAEP Achievement Levels

CONTINUED FROM PREVIOUS PAGE

"work was immediately stopped and its report never officially released."

Validity of Scores

The validity of the information derived from NAEP scores also has come under question. Writing in the Fall 1991 *Educational Measurement*, Robert Forsyth argues that the NAEP proficiency scales in math and science do not yield "valid criterion-referenced interpretations." He charged that NAGB has been drawing conclusions from the test scores that cannot be supported by the data. His study lends support to the charges that the achievement levels set by NAGB are inadequate.

Additionally, Forsyth questioned whether students who take the NAEP tests care enough about them to try to do well. Speaking before the National Association of State Boards of Education annual meeting, NAGB Chair Richard Boyd reported that the NAEP results may underestimate actual capabilities for this reason.

✎ "Summative Evaluation of the NAGB's Inaugural 1990-1991 Effort to Set Achievement Levels on the NAEP" and NAGB's response to the Evaluation are available by requesting them under the Freedom of Information Act. Write NAGB, 1100 L St. NW, #7322, Washington, DC 20005-4013; or fax (202) 357-6945. The Linn report is now available from CRESST, CSE Report #330 ($6); call Kim Hurst, (213) 206-1512. ❑

Panel Urges Indicator System

A Special Study Panel on Education Indicators has recommended development of a new national system for collecting information about the status of U.S. schools. The panel, mandated by Congress in 1988, called for an indicator system to be organized around six "enduring educational issues:" learner outcomes, quality of educational institutions, readiness for school, societal support for learning, education and economic productivity, and equity.

"The panel's conception of learner outcomes goes far beyond testing student achievement in a limited number of subjects," the panel wrote in its report, *Education Counts*, adding, "The nation needs a much broader definition of appropriate learner outcomes."

✎ *Education Counts* is available for $6.50 from Superintendent of Documents, US Government Printing Office, Washington, DC 20402-9325; stock number 065-000-00463-1. ❑

NEGB, NAGB, and Dept. of Ed. Release Reports on Test Results

On September 30, the National Educational Goals Panel (NEGB) released its first "Report Card" on US education, the National Assessment Governing Board (NAGB) issued its report on student performance on the 1990 math exam in terms of specific achievement levels, and the Department of Education released data from a National Assessment of Educational Progress (NAEP) study examining trends in student achievement over the past twenty years.

Information on the "Report Card," as well as reports of the technical planning subgroups for each goal, are available from NEGB, 1850 M St., NW, #270, Washington, DC 20036; 202-632-0952.

"The LEVELS of Mathematics Achievement" reports on student performance at grades 4, 8 and 12 on the 1990 NAEP math exam (see "Controversy Erupts over NAEP"). In addition to "Volume I, National and State Summaries," and "Volume II, State Results for Released Items" (grade 8), reports are available on each state that participated in the trial state assessments. Contact NAGB, 1100 L St, NW, #7322, Washington, DC 20005-4013.

The full NAEP trend report, which found that US students were achieving at about the same levels as they were in 1970, has not yet been released. ❑

WORTH READING

Rethinking Assessment

Two recent works that make significant contributions to the fundamental rethinking of assessment and measurement are John Raven, *The Tragic Illusion: Educational Testing* (Trillium Press, Uniondale, NY) and Dennie Wolf, Janet Bixby, John Glenn III and Howard Gardner, "To Use Their Minds Well: Investigating New Forms of Student Assessment" (in G. Grant, ed., *Review of Research in Education 17*, American Educational Review Association, Washington, DC). Both offer powerful critiques of traditional measurement theory and practice, and new directions for assessment. ❑

Connecticut's Performance Assessments

Connecticut has entered the third year of development of math and science performance assessments for its Common Core of Learning, which outlines the skills and knowledge expected of the state's high school students.

With a $1 million grant from the National Science Foundation, Connecticut began developing the assessments in 1989 in collaboration with Michigan, Minnesota, Texas, Vermont, New York, Wisconsin, and members of the Coalition of Essential Schools.

In the first year, assessment tasks were designed primarily by teachers, with the assistance of scientists, mathematicians and other project consultants. Teachers involved in the design and piloting of "task prototypes" were volunteers nominated by each state consultant. A task prototype is a model assessment activity done by students which can serve as the basis for development of other assessment tasks within the classroom.

Some of the approximately 150 teachers who have participated in the project are from schools in some of the largest cities in the country. This increases the potential for addressing the issues of equity and generalizability of the assessments.

During the second year of the project, nearly fifty tasks were tried out in classrooms. The tasks are embedded in the curriculum and involve groups of students designing and carrying out investigations which include data gathering, problem-solving and oral presentations. Tasks that are "embedded in the curriculum" are integrated into the school year by teachers at a time determined by them.

A Homework Helper for Elenore

You are babysitting Elenore, who has just finished her arithmetic homework. "I really like doing decimals! It's easy," the youngster boasts as she proudly shows you her work:

$$\begin{array}{r} .8 \\ +\ .4 \\ \hline .12 \end{array} \qquad \begin{array}{r} .3 \\ +\ .6 \\ \hline .9 \end{array} \qquad \begin{array}{r} .7 \\ +\ .9 \\ \hline .16 \end{array}$$

Is Elenore's homework correct? If not, write what you would tell and/or show Elenore, to help her *understand* her homework.

Tasks can last from a few days to a week or more. As students work on the tasks, they are both learning and being assessed. Ideally, teachers will use a combination of model tasks and ones which they develop for their own classrooms.

According to Joan Baron, Coordinator of the Common Core of Learning Assessment, among the difficulties encountered during the development of the assessments were the absence of a "vocabulary" for describing and discussing the kinds of tasks which elicit demonstrations of what students know and are able to do, and the lack of curricula which enables these processes to occur.

Despite the difficulties involved in designing and implementing assessments, teachers report being empowered by the work they are doing. They have reported that students respond positively to the tasks, challenging though they have been. Group work especially has proven to be a powerful experience for students and teachers, explained Baron.

Teachers also have received training to help them ease students into cooperative learning activities. They have learned to determine what students know before undertaking the group assessment tasks and to evaluate whether working in groups improves student learning.

Connecticut's program calls for summative essay questions in math and science in grade 11 which are open-ended and allow for many different solutions. According to Baron, these Connecticut assessments will not have "high stakes" for students. The assessments are for program accountability and instructional improvement purposes. Because the material is new, students will not be held responsible for what they have not had the opportunity to learn. Once the assessments are fully implemented, the state will annually evaluate the performance of a random sample of eleventh graders against the objectives of the Common Core of Learning.

During the 1991-92 school year, project participants will work on further developing task prototypes and scoring guides, including procedures for how to aggregate scores across districts.

✎ For more information, Connecticut Common Core of Learning Assessment Project, Room 340, Connecticut Department of Education, Box 2219, Hartford, CT 06145. ❑

Controversial Oregon Reforms Pose Questions for Future of Assessment

Oregon has become the first state to adopt "education reform" legislation modeled on proposals from the National Center on Education and the Economy (NCEE). The legislation includes plans for a new assessment system.

Many education reformers fear the new law will lead to more, not better, testing, increased tracking, and decreased equity.

State Rep. Vera Katz, who spearheaded the Oregon law's adoption, was a NCEE Board member. The NCEE, together with the Learning Research and Development Center, has created the New Standards Project, which is developing model exams for a national examination system. Thus, the Oregon legislation provides a preview of what might happen if a national exam system is ever implemented. Federal legislation modeled on the NCEE proposals has been introduced, and Virginia's Gov. Wilder has proposed similar legislation.

Oregon's H.B.3565 is meant to restructure the state's schools, not just introduce new testing. Its centerpiece is a program for "professional and technical training" of students — labelled "vocational training" by critics. The law also includes initiatives for early childhood education, non-graded primary programs, site-based management, remedial education, drop-out prevention, and more.

Supporters and critics agree that the ultimate outcome of the legislation will depend on how the Oregon Department of Education (ODE) writes program rules and regulations, the way the programs are implemented by districts, and the availability of additional funding. Because of a fiscal crisis in Oregon, no money has been appropriated to pursue these programs.

In the absence of funding and clear program definition in the law, critics fear that implementation will mean intensified testing and tracking and the elimination of the last two years of high school for low-income youth.

Testing

The law mandates testing "at approximately grades 3, 5, 8, and 10 including but not limited to work samples, tests, and portfolios. Students shall be allowed to collect credentials over a period of years, culminating in a project or exhibition that demonstrates attainment of the required knowledge and skills." Once reaching this level, at about grade 10, the student receives a Certificate of Initial Mastery (CIM, also called for in NCEE planning). Students would then become eligible to earn a Certificate of Advanced Mastery (CAM).

The law requires that the test scores of individuals be used as part of school and state report cards. However, use of performance assessments to report individual, school and district information for many grades in a number of subjects has never been done anywhere in the country.

Currently, Oregon administers multiple-choice tests in grades 3, 5, 8 and 10. The law calls for continuing these tests until the new plan "is operational." The bill also says that by 1997, "every student shall have the opportunity...to obtain" a CIM. The law is not clear whether the CIM's would be awarded on the basis of scores on multiple-choice tests if performance assessments have not been implemented. No financing exists for transforming the system to one of performance exams, portfolios and projects, or for the necessary staff development.

Tracking?

On attainment of the CIM, students choosing to continue education in what would have been their junior and senior years of high school must select a program leading to a CAM "and a college preparatory or academic professional technical indorsement, or both." The legislation says the State Board of Education "shall facilitate the movement between the indorsements." Critics argue that the program will lead to tracking before and after the CIM, and ask why, if the two are comparable, two post-CIM tracks are needed.

Although the legislation calls for "instructional groupings" to be organized "as heterogeneously as possible," the Committee to Revise the Katz Education Plan says that the law, "by requiring two classes of students in the high schools, legitimates and will intensify tracking throughout the school system." Moreover, the vocational track will, in their view, remain seriously underfunded.

CONTINUED ON NEXT PAGE

Oregon Reforms Pose Questions for Future

CONTINUED FROM PREVIOUS PAGE

Equity

The new law also states that Oregon "(p)rovides equal and open access and educational opportunities for all students in the state." However, it does nothing to equalize funding among districts or programs. The FairTest-initiated Campaign for Genuine Accountability in Education has argued that no national examination system should be implemented unless an equal opportunity to pass the tests is ensured. Similarly, Oregon critics are concerned that all students will not have an equal opportunity to do well on the tests for the CIM. As a result, many will be tracked into dead-end vocational programs or will drop out of school. This could allow the state to end its responsibility for educating many low-income 17- and 18-year-olds.

Reforming the Reform

Critics of the bill will be working for legislative amendments, favorable ODE rules and guidelines, and sensitive implementation by the districts.
In particular, the Committee to Revise the Katz Education Plan has called for a reduction in the amount of testing. "(T)he potentials for high pressure testing should be eliminated," said the Committee.

✎ For more information, contact the Committee to Revise the Katz Education Plan, 1615 SE 35th Place, Portland, OR 97214. ☐

More Hurdles for Prospective Teachers

CONTINUED FROM PAGE SIXTEEN

accurately predict classroom performance (see *Examiner*, Fall, 1988).

Civil rights and education reform groups have pointed to the disproportionate failure rates of Blacks, Hispanics and women on the NTE and other ETS teacher tests. Many advocates believe the tests have played a major role in keeping minority educators out of classrooms and even deterred many from entering teacher training programs (see *Examiner*, Winter, 1989). As a result, several states have recently considered legislation dropping their NTE requirement, and Rhode Island adopted such a policy this year.

To date, there is no proof that ETS has addressed these significant concerns in designing the Praxis Series of tests. Skeptics believe the new tests respond more to the test-maker's growing public relations problems than to the need for fair, accurate and relevant assessment of entering teachers. They will urge state policy-makers to carefully examine new tests and determine their likely impact before mandating them for any students. ☐

Rhode Island Drops NTE Requirement

Despite intensive lobbying efforts by the Educational Testing Service (ETS), Rhode Island has adopted a new law dropping its requirement that all applicants for state teaching licenses pass the National Teachers Examination (NTE). The provision will also apply to the new series of teacher exams being developed by ETS.

The landmark statute, which culminates a five-year campaign led by the Rhode Island Affiliate of the American Civil Liberties Union (ACLU), states, "No person shall be barred from certification under this chapter solely because of his or her score on any standardized teachers examination." The state Board of Regents may require anyone who does not pass such a test to submit to an annual "job performance review" for three years.

According to Rhode Island ACLU Executive Director Steven Brown, who coordinated the successful campaign, ETS hired a former Rhode Island legislator to lobby against the bill, but his attempts to scuttle it failed. Advocates for the new policy, including FairTest, argued that reliance on a standardized test to determine who would make a good teacher made no sense and that the NTE had demonstrable cultural biases.

Brown concluded, "We hope that passage of the law here will encourage other states to adopt similar legislation. The discriminatory nature of teacher testing is not confined to Rhode Island. We want this to be a national model."

Special credit goes to Brown, Rep. Harold Metts who filed the original NTE moratorium bill, Sen. Charles Walton who persuaded his peers to pass the final version, and the Urban League of Rhode Island for their effective grassroots pressure.

✎ For more information, contact Steven Brown, RI-ACLU, 212 Union St., Providence, RI 02903. ☐

☐ **FAIRTEST EXAMINER** $________
published quarterly…$15/year (individuals), $25/year (institutions). Back issues $4: Vol.____ No.____

☐ **THE SAT COACHING COVER-UP**, *by FairTest staff* (see page 1) …$7.95 $________

☐ **STANDING UP TO THE SAT**, *by John Weiss, Barbara Beckwith and Bob Schaeffer*, 192 pp.…$6.95 $________

☐ **FALLOUT FROM THE TESTING EXPLOSION: How 100 Million Standardized Exams Undermine Equity and Excellence in America's Public Schools**, *by Noe Medina and Monty Neill*, 3rd edition (77 pp.)… $8.95 $________

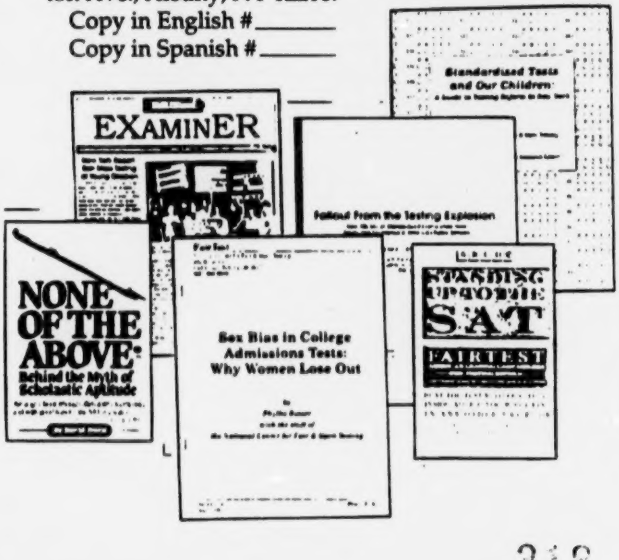
☐ **SEX BIAS IN COLLEGE ADMISSIONS TESTS: WHY WOMEN LOSE OUT**, *by Phyllis Rosser and FairTest staff*, 3rd. edition (60 pp)…$7.95 $________

☐ **THE REIGN OF ETS**, *by Allan Nairn and Ralph Nader* (550 pp)…$30.00 $________

☐ **BEYOND STANDARDIZED TESTS: ADMISSIONS ALTERNATIVES THAT WORK**, *by Amy Allina with FairTest staff* (18 pp)…$5.50 $________

☐ **STANDARDIZED TESTS AND OUR CHILDREN: A GUIDE TO TESTING REFORM**, *by FairTest* (32 pp.)…$4; 5 for $15; 10 for $20; 50 for $50; 100 for $75. English and Spanish editions can be combined for bulk prices. $________
Copies in English #______
Copies in Spanish #______

☐ **STANDARDIZED TESTS AND OUR CHILDREN: NEW YORK EDITION**, *by FairTest & NYPIRC* (32 pp.) same prices as above. Single copies from FairTest; bulk copies from NYPIRC, 184 Washington Ave., Albany, NY 12210. $________
Copy in English #______
Copy in Spanish #______


EXAMINER
Standardized Tests and Our Children
Fallout From the Testing Explosion
NONE OF THE ABOVE: Behind the Myth of Scholastic Aptitude
Sex Bias in College Admissions Tests: Why Women Lose Out
STANDING UP TO THE SAT
FAIRTEST

CONTRIBUTIONS

I agree there's no better way to be part of the testing reform movement than supporting FairTest! Enclosed is my tax-deductible contribution.

☐ **FAIRTEST ASSOCIATE**
$25—includes a free subscription to the FairTest *Examiner*

☐ **FAIRTEST SPONSOR**
$50—includes Examiner subscription and a FairTest publication of your choice, excluding Reign of ETS.

☐ **FAIRTEST SUSTAINER**
$100 or more—includes above, plus a 20% discount on publications.

☐ **OTHER**
____
____

TOTALS

PUBLICATIONS TOTAL
$ ____________

CONTRIBUTIONS TOTAL
$ ____________

CHECK TOTAL
$ ____________

Make payable to FairTest and mail to 342 Broadway, Cambridge, MA 02139. All contributions to FairTest are fully tax deductible, except for the value of publications received.

NAME ________________________
ADDRESS ________________________
CITY ________________________
STATE ____________ ZIP ________

EMPLOYMENT TESTING

The "All New, Improved NTE": More Hurdles for Prospective Teachers

The Educational Testing Service has announced that it will soon replace its much-criticized National Teachers Exam (NTE) with a series of three tests it hopes all prospective teachers will be required to take. The "new generation" of tests bears the high-sounding title of "The Praxis Series: Professional Assessments for Beginning Teachers" but appears to be little more than a repackaging of current ETS exams administered to prospective educators.

Costs for the new battery of exams have not yet been made public, but it is likely that students will pay substantially more than the $80 now charged for the NTE Core Battery. Thirty-three states and the District of Columbia currently use the NTE or one of several other ETS-designed teacher tests such as the Pre-Professional Skills Test (PPST) and California's CBEST.

The first Praxis Series tests are scheduled to be administered in the fall of 1992 with regular use beginning in the 1993-94 school year. The three exams, which ETS calls "Stages," include:

1) A computer-administered "basic skills" exam in reading, writing and mathematics which typically will be taken during a college student's sophomore year. Though one essay question will be included, it will be predominantly multiple-choice.

2) A battery of subject-area assessments, including a two-hour common core of multiple-choice questions and optional subject-specific exams based on constructed responses in about a dozen areas. Teacher candidates are expected to take these tests when they are ready to enter the job market after completing their undergraduate (or graduate) programs. This stage consolidates parts of the NTE Core Battery and Specialty Area Tests.

3) Yet another assessment will be administered during the first year of classroom teaching. Like "vaporware" computer programs that are heavily publicized before they actually have been written, few details are available about this final, "performance-based" hurdle for prospective teachers which may be the only truly new part of the testing series.

Over the past decade, the NTE and other ETS teacher tests have received an increasing number of negative reviews. Academic experts have criticized the NTE's content (one study found that 40% of items lacked a single correct answer), multiple-choice format, and pedagogic assumptions. Most importantly, they have argued that there is no evidence the NTE can

CONTINUED ON PAGE 14

FairTest
National Center for Fair & Open Testing
342 Broadway
Cambridge, MA 02139-1802
(617) 864-4810
(617) 497-2224 (fax)

Non-Profit
U.S. Postage
PAID
Cambridge, MA
Permit No. 53278

Address Correction Requested

Do you have an asterisk (*) on your label?
If you do, this is your last issue. **Renew now!**

Moving? Return your label with new address